

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 24, 2005, 06:59:05 ; Search time 180 seconds  
(without alignments)

1300.112 Million cell updates/sec

Title: US-10-691-374-2

Perfect score: 2390

Sequence: 1 NWI09NSILLIILIFASSIS.....VTPHCTSLSEIDELVNY 457

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :  
1: uniprot\_sprot:\*  
2: uniprot\_trembl:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score  | Query Match | Length | DB ID | Description        |
|------------|--------|-------------|--------|-------|--------------------|
| 1          | 2390   | 100.0       | 457    | 1     | P05117 lycopersico |
| 2          | 1295.5 | 54.2        | 467    | 1     | PGLR_LYCCH         |
| 3          | 1282   | 53.6        | 453    | 2     | O9M6S2             |
| 4          | 1204.5 | 50.4        | 444    | 2     | O94B15             |
| 5          | 1185.5 | 49.6        | 428    | 2     | O944B5             |
| 6          | 1185   | 49.6        | 438    | 2     | O9SEB7             |
| 7          | 1149   | 48.1        | 433    | 2     | O8RY29             |
| 8          | 1136   | 47.5        | 462    | 1     | PGLR_PERRAE        |
| 9          | 1130   | 47.3        | 434    | 2     | O9FDT7             |
| 10         | 1128   | 47.2        | 461    | 2     | O81246             |
| 11         | 1127   | 47.2        | 431    | 2     | O23147             |
| 12         | 1115.5 | 46.7        | 426    | 2     | O22935             |
| 13         | 1113   | 46.6        | 433    | 2     | O42399             |
| 14         | 1108   | 46.4        | 458    | 2     | O43063             |
| 15         | 1106   | 46.3        | 433    | 2     | O8VXT3             |
| 16         | 1103   | 46.2        | 433    | 2     | O42636             |
| 17         | 1091   | 45.6        | 460    | 1     | P48978 MALDO       |
| 18         | 1073   | 44.9        | 460    | 2     | O8GTF8             |
| 19         | 1061.5 | 44.4        | 335    | 2     | O65886             |
| 20         | 1059   | 44.3        | 199    | 2     | O70V32             |
| 21         | 1043   | 43.6        | 460    | 2     | O68VK1             |
| 22         | 1034.5 | 43.3        | 405    | 2     | O84ZP3             |
| 23         | 1022.5 | 42.8        | 422    | 2     | O84ZP5             |
| 24         | 1014   | 42.4        | 429    | 2     | O84ZP9             |
| 25         | 998.5  | 41.8        | 200    | 2     | O70V18             |
| 26         | 988.5  | 41.4        | 444    | 2     | O84W40             |
| 27         | 925    | 38.7        | 452    | 2     | O9MTD3             |
| 28         | 914    | 38.2        | 444    | 2     | O94AD5             |
| 29         | 898.5  | 37.6        | 459    | 2     | O9SSC2             |
| 30         | 840.5  | 35.2        | 468    | 2     | O9CAL5             |
| 31         | 825.5  | 34.5        | 435    | 2     | O9SLP3             |

|    |       |      |     |   |            |                    |
|----|-------|------|-----|---|------------|--------------------|
| 32 | 807   | 33.8 | 481 | 2 | O6ZGAI     | O6ZGAI oryza sativ |
| 33 | 802   | 33.6 | 426 | 2 | O65XN6     | O65XN6 oryza sativ |
| 34 | 796.5 | 33.3 | 514 | 1 | PGL2_CHAOB | O7M1E7 chamecypr   |
| 35 | 794   | 33.2 | 408 | 2 | O65XN5     | O65XN5 oryza sativ |
| 36 | 789   | 33.0 | 514 | 2 | O8H987     | O8H987 cryptomeria |
| 37 | 788.5 | 33.0 | 514 | 2 | O8H989     | O8H989 cryptomeria |
| 38 | 788   | 33.0 | 457 | 2 | O9ZUE7     | O9ZUE7 arabidopsis |
| 39 | 787.5 | 32.9 | 434 | 2 | O9FXC1     | O9FXC1 arabidopsis |
| 40 | 787.5 | 32.9 | 514 | 1 | PGL2_CRYUA | P4312 cryptomeria  |
| 41 | 787.5 | 32.9 | 514 | 2 | O8H988     | O8H988 cryptomeria |
| 42 | 764   | 32.0 | 507 | 1 | PGL2_TUNAS | O9FY19 juniperus a |
| 43 | 749.5 | 31.4 | 161 | 2 | O9LQD1     | O9LQD1 arabidopsis |
| 44 | 741.5 | 31.0 | 491 | 2 | O9FWX5     | O9FWX5 arabidopsis |
| 45 | 736   | 30.8 | 423 | 2 | O81245     | O81245 cucumis mel |

#### ALIGNMENTS

RESULT 1  
PGLR\_LYCCHES STANDARD; PRT; 457 AA.  
ID PGLR\_LYCCHES  
AC P05117;  
DT 13-AUG-1987 (Rel. 05, Created)  
DT 13-AUG-1987 (Rel. 05, Last sequence update)  
DE 25-OCT-2004 (Rel. 45, Last annotation update)  
DE Polysaccharonase 2A precursor (EC 3.2.1.15) (PG-2A) (Pectinase).  
OS Lycopersicon esculentum (Tomato).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;  
OC Lamiales; Solanales; Solanaceae; Solanum.  
OX NCBI\_TaxID=4081;  
RN (1)  
RP SEQUENCE FROM N.A.  
RC STRAIN=cv. Alisa Craig;  
RA Bird C.R., Smith C.J.S., Ray J.A., Mourreau P., Bevan M.W., Bird A.S.,  
RA Hughes S., Morris P.C., Grierson D., Schuch W.;  
RT "The tomato polygalacturonase gene and ripening-specific expression in  
transgenic plants";  
RL Plant Mol. Biol. 11:651-662(1988).  
RN (2)  
RP SEQUENCE FROM N.A.  
RC STRAIN=cv. Alisa Craig;  
RX MEDLINE=87066731; PubMed=3786135;  
RA Grierson D., Tucker G.A., Keen J., Ray J., Bird C.R., Schuch W.;  
RT "Sequencing and identification of a cDNA clone for tomato  
polygalacturonase.";  
RL Nucleic Acids Res. 14:8595-8603(1986).  
RN (3)  
RP SEQUENCE FROM N.A.  
RA Sheehy R.E., Pearson J., Brady C.J., Hlatk W.R.;  
RT "Molecular characterization of tomato fruit polygalacturonase.";  
RL Mol. Gen. Genet. 208:30-36(1987).  
RN (4)  
RP REVISIONS.  
RA Hlatk W.R.;  
RT Submitted (OCT-1987) to the EMBL/GenBank/DBJ databases.  
CC -!- FUNCTION: Acts in concert with the pectinesterase, in the ripening  
process. Is involved in cell wall metabolism, specifically in  
polyuronide degradation.  
CC CATALYTIC ACTIVITY: Random hydrolysis of 1,4-alpha-D-  
galactosiduronic linkages in pectate and other galacturonans.  
CC SUBCELLULAR LOCATION: Secreted.  
CC -!- DEVELOPMENTAL STAGE: In ripening fruit.  
CC -!- BIOTECHNOLOGY: The effect of PG can be neutralized by introducing  
an antisense PG gene by genetic manipulation. The Flavr Savr  
tomato produced by Calgene (Monsanto) in such a manner has a  
longer shelf life due to delayed ripening.  
CC -!- SIMILARITY: Belongs to the glycosyl hydrolase 28 family.  
CC -!- SIMILARITY: Contains 4 Pdh1 repeats.  
CC This Swiss-Prot entry is copyrighted. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -

the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).

DR EMBL; X14074; CA32235.1; -  
 DR EMBL; M37304; AAA34178.1; -  
 DR EMBL; X04583; CA28254.1; -  
 DR EMBL; X05656; CA29148.1; -  
 DR EMBL; A15981; CA01256.1; -  
 DR EMBL; A24194; CA01720.1; -  
 DR PIR; A25534; A25534; -  
 DR InterPro; IPR000743; Glyco\_hydro\_28.  
 DR InterPro; IPR006626; PBH1.  
 DR InterPro; IPR011050; Pectin\_lyase\_1like.  
 DR Pfam; PF00295; Glyco\_hydro\_28; 1.  
 DR SMART; SM00710; PBH1; 4.  
 DR PROSITE; PS00502; POLYGALECTURONASE; 1.  
 DR Cell wall; Fruit ripening; Genetically modified food; Glycoprotein;  
 KM Glycosidase; Hydrolase; Repeat; Signal.  
 FT SIGNAL 1 24 Potential.  
 FT PROPEP 25 71  
 FT CHAIN 72 457 Polygalacturonase 2A.  
 FT REPEAT 228 255 PBH1 1.  
 FT REPEAT 256 277 PBH1 2.  
 FT REPEAT 309 330 PBH1 3.  
 FT REPEAT 338 359 PBH1 4.  
 FT ACT SITE 293 293 Probable.  
 FT CARBOHYD 189 189 N-linked (GlcNAc...) (Potential).  
 FT CARBOHYD 240 240 N-linked (GlcNAc...) (Potential).  
 FT CARBOHYD 286 286 N-linked (GlcNAc...) (Potential).  
 FT CARBOHYD 311 311 N-linked (GlcNAc...) (Potential).  
 SQ SEQUENCE 457 AA; 50051 MW; 449E4DC369198074 CRC64;

Query Match  
 Best Local Similarity 100.0%; Score 2390; DB 1; Length 457;  
 Matches 457; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVIQNSILLIIIFASSISTCRSNVIDNLFKQVYNNIIIEOFADHFOAYLSYSKNT 60  
 DB 1 MVIQNSILLIIIFASSISTCRSNVIDNLFKQVYNNIIIEOFADHFOAYLSYSKNT 60  
 QY 61 SNNNIDKVDKNGIVINVLSPGAGGDKTYDNIAFEQANWEACSSRTPVQFVVKKNY 120  
 DB 61 SNNNIDKVDKNGIVINVLSPGAGGDKTYDNIAFEQANWEACSSRTPVQFVVKKNY 120  
 QY 121 LKQITFGSPCRSSISVKIFGSLKASRKISDYKDRRLMIAFDSVQNLVGGGGTNGGV 180  
 DB 121 LKQITFGSPCRSSISVKIFGSLKASRKISDYKDRRLMIAFDSVQNLVGGGGTNGGV 180  
 QY 121 LKQITFGSPCRSSISVKIFGSLKASRKISDYKDRRLMIAFDSVQNLVGGGGTNGGV 180  
 DB 121 LKQITFGSPCRSSISVKIFGSLKASRKISDYKDRRLMIAFDSVQNLVGGGGTNGGV 180  
 QY 181 WMPSSCKINSLPERDAPFLALTFWNCNKLKVNLLKSKNAQOIHKFESCTNVVASLMTN 240  
 DB 181 WMPSSCKINSLPERDAPFLALTFWNCNKLKVNLLKSKNAQOIHKFESCTNVVASLMTN 240  
 QY 241 ASAKSPNTDGVHVSNTQYIOISPTIIGTDDCISVSGSQNVQANTITGPHGISISL 300  
 DB 241 ASAKSPNTDGVHVSNTQYIOISPTIIGTDDCISVSGSQNVQANTITGPHGISISL 300  
 QY 301 GSNSNAYVSNVTNVEAKIIGAENGVRKTMWGGSGQASNIKFLVNEQDVKYPIIIDN 360  
 DB 301 GSNSNAYVSNVTNVEAKIIGAENGVRKTMWGGSGQASNIKFLVNEQDVKYPIIIDN 360  
 QY 361 YCDRVEPCIQPSAVQNVVYENKGNATYVAIKFPCSTNPFCEGIMENINLVESG 420  
 DB 361 YCDRVEPCIQPSAVQNVVYENKGNATYVAIKFPCSTNPFCEGIMENINLVESG 420  
 QY 421 KRSKATCKNVHNNAEHTVPHCTSLSEISDEBALYNY 457  
 DB 421 KRSKATCKNVHNNAEHTVPHCTSLSEISDEBALYNY 457

RESULT 2

PGRL ACTCH ID PGRL ACTCH STANDARD; PRT; 467 AA.  
 AC P35316;  
 DT 01-JUN-1994 (Rel. 29, Created)  
 DT 01-JUN-1994 (Rel. 29, Last sequence update)  
 DT 25-OCT-2004 (Rel. 45, Last annotation update)  
 DE Polygalacturonase precursor (EC 3.2.1.15) (PG) (Pectinase).  
 OS Actinidia chinensis (Kiwifruit) (Yangtze).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;  
 OC Ericales; Actinidiaceae; Actinidia.  
 OX NCBI\_TaxID=3625;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Deliciosa; Pubmed=8029342; DOI=10.1104/pp.103.2.669;  
 RX MEDLINE=94302157; Pubmed=8029342; DOI=10.1104/pp.103.2.669;  
 RA Atkinson R.G., Gardner R.C.;  
 RA "A polygalacturonase gene from kiwifruit (Actinidia deliciosa).",  
 RL Plant Physiol. 103:669-670 (1993).  
 CC -1- FUNCTION: Acts in concert with the pectinesterase, in the ripening process. Is involved in cell wall metabolism, specifically in polyuronide degradation.  
 CC -1- CATALYTIC ACTIVITY: Random hydrolysis of 1,4-alpha-D-galactosiduronic linkages in pectate and other galacturonans.  
 CC -1- SUBCELLULAR LOCATION: Secreted.  
 CC -1- DEVELOPMENTAL STAGE: In ripening fruit.  
 CC -1- SIMILARITY: Belongs to the glycoyl hydrolase 28 family.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).

DR EMBL; L12019; AAC14453.1; -  
 DR InterPro; IPR000743; Glyco\_hydro\_28.  
 DR InterPro; IPR006626; PBH1.  
 DR InterPro; IPR011050; Pectin\_lyase\_1like.  
 DR Pfam; PF00295; Glyco\_hydro\_28; 1.  
 DR SMART; SM00710; PBH1; 4.  
 DR PROSITE; PS00502; POLYGALECTURONASE; 1.  
 DR Cell wall; Fruit ripening; Glycoprotein; Glycosidase; Hydrolase;  
 KM Signal.  
 FT SIGNAL 1 27 Potential.  
 FT CHAIN 28 467 Polygalacturonase.  
 FT ACT SITE 306 306 Probable.  
 FT CARBOHYD 290 290 N-linked (GlcNAc...) (Potential).  
 SQ SEQUENCE 467 AA; 50776 MW; 5A9A61483C028B7A CRC64;

Query Match  
 Best Local Similarity 54.2%; Score 1295.5; DB 1; Length 467;  
 Matches 257; Conservative 75; Mismatches 112; Indels 27; Gaps 7;

QY 1 MVIQNR--SILLIIIFASSISTCRSNVIDNLFKQVYNNIIIEOFADHFOAYLSYSKNT 58  
 DB 1 MVIQNR--SILLIIIFASSISTCRSNVIDNLFKQVYNNIIIEOFADHFOAYLSYSKNT 58  
 QY 59 IESNNNIDKVDKNGI-----KYNVLSPGAGGDKTYDNIAFEQANWE 101  
 DB 59 IESNNNIDKVDKNGI-----KYNVLSPGAGGDKTYDNIAFEQANWE 101  
 QY 102 ACSSRTPVQFVVKKNYLLKQITFGSPCRSSISVKIFGSLKASRKISDY-KDRRLMIAF 160  
 DB 102 ACSSRTPVQFVVKKNYLLKQITFGSPCRSSISVKIFGSLKASRKISDY-KDRRLMIAF 160  
 QY 115 ACSSTSSAVLLVLPK-KNVLVPIISFGSPCKSGLMWYGTITKASDDRSVDKRGKRLVLP 173  
 DB 115 ACSSTSSAVLLVLPK-KNVLVPIISFGSPCKSGLMWYGTITKASDDRSVDKRGKRLVLP 173  
 QY 161 DSVQNLVVGSGGTNGGVWMPSSCKINSLPERDAPFLALTFWNCNKLKVNLLKSKNAQ 220  
 DB 161 DSVQNLVVGSGGTNGGVWMPSSCKINSLPERDAPFLALTFWNCNKLKVNLLKSKNAQ 220  
 QY 221 QIHKFESCTNVVASLMTNIAKSPNTDGVHVSNTQYIOISPTIIGTDDCISVSGSQ 280  
 DB 221 QIHKFESCTNVVASLMTNIAKSPNTDGVHVSNTQYIOISPTIIGTDDCISVSGSQ 280

Db 234 QIHVSFDCNVVQASNLMTAPENSPTDGIHVTGQNIHSSCVIGTGDICISIVNGSR 293  
Qy 281 NVOATNTGCPGHGISIGSGNSSEAYVSNVTVNEAKIIGAENGVRKIKTWGGSGQASN 340  
Db 294 KAVVNDITGCPGHGISIGSGNSSEAYVSNVTVNEAKIIGAENGVRKIKTWGGSGQASN 353  
Qy 341 IKFLANEMODVKYPIIIDQNYCDRVEPCIQFSAVQVKNVYENIKGTSATKYAIKEDCS 400  
Db 354 IKFQNVEMHNVENPIIIDQNYCDQDKPCQEGSSAVQVKNVYQVNIKGTCA SNVAITFDGS 413  
Qy 401 TNFPCGIIEMNTLVGSGKPEATCKVNFNNAEHTPHCTSLSEDE 451  
Db 414 KRFPCGIIVEDVLEIEGGAALKCNVELSETGVSPHCEEGEE 464

## RESULT 3

Qy 09M6S2 PRELIMINARY; PRT; 463 AA.  
ID Q9M6S2  
AC Q9M6S2  
DT 01-OCT-2000 (TEMBLrel. 15, Created)  
DT 01-OCT-2000 (TEMBLrel. 15, Last sequence update)  
DT 01-MAR-2004 (TEMBLrel. 26, Last annotation update)  
DE Polyalacturonase A.  
GN NameA6A;  
OS Actinidia chinensis (Kiwi) (Yangtze).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;  
OC Ericales; Actinidiales; Actinidia.  
OX NCBI\_TaxId=3625;  
RN (1)  
RP SEQUENCE FROM N.A.  
RC TISSUE=Fruit pericarp;  
RX MEDLINE=20252524; PubMed=10794531; DOI=10.1006/30952922;  
RA Wang Z.Y., Macrae E.A., Wright M.A., Bolitho K.M., Rose G.S.,  
RA Atkinson R.G.;  
RT "Polygalacturonase gene expression in kiwifruit: relationship to fruit  
RT aftertong and ethylene production";  
RL Plant Mol. Biol. 42:317-328(2000).  
CC -1- SIMILARITY: Belongs to family 28 of glycosyl hydrolases.  
DR EMBL: AF152578; XAF71160.1; -;  
DR GO: GO:000518; C:cell wall; IEA.  
DR GO: GO:0016798; F:hydrolase activity, acting on glycosyl bonds; IEA.  
DR GO: GO:0004650; F:polygalacturonase activity; IEA.  
DR GO: GO:0005975; P:carbohydrate metabolism; IEA.  
DR InterPro: IPR000743; Glyco\_hydro\_28.  
DR InterPro: IPR006626; Pbh1.  
DR InterPro: IPR011050; Pectin\_lyase\_like.  
DR InterPro: IPR000408; Reg\_chit\_condens.  
DR Pfam: PF00295; Glyco\_hydro\_28; 1.  
DR SMART: SM00710; Pbh1; 4.  
DR PROSITE: PS00502; POLYGALACTURONASE; 1.  
DR PROSITE: PS00626; RCCL\_2; UNKNOWN\_1.  
DR Cell wall; Glycosidase; Hydrolase.  
KW SEQUENCE 463 AA; 50352 MW; 7EEDFEFE54988FA CRC64;

Query Match 53.6%; Score 1282; DB 2; Length 463;  
Best Local Similarity 54.2%; Pred. No. 1.8e-81;  
Matches 256; Conservative 72; Mismatches 114; Indels 30; Gaps 8;

Qy 1 MVIQNR--SILLIIPASSISTCRSNVIDNLFKQYVDNILEQEFANDEFOAYLSYLSKN 58  
Db 1 MAQRHNFQFVITLILPSFILGTSVNHED---PRHDVHLE-EYGVDPKAYPSYITTI 55  
Qy 59 IESNNNNIDKVDKNGI-----KVINLVSGANGDGKTYNIAEQAMNE 101  
Db 56 GDDNDFGSGMSHENGIFGLRKVDYGMDFVLDPASTVNVNDDFGAGKDGSD-DTKAFKAKWA 114  
Qy 102 AGSRRTPVQGVQPKNKVYLKQITFGSPCRSISVKIFGLSEASKSISDY-KDRRLMIAP 160  
Db 115 VCSSTSAVLVLVPO-KNYLVRLPIFGSPCSDLTMQIVGTLEASDDNSDKGRHMLVF 173  
Qy 161 DSVQNLVVGCGGTNGGVWMPSSCKINKSLPCRDAPITALTFWNCIGLVNNLKSQAO 220

Db 174 DSVQNLVVBGGGTNGNGQIMQNSCKNTKLPCKDAPITALFFYKSKHVI VKNLKIENAO 223  
Qy 221 QIHIFESCTNVVNASNLMTINASAKSPNTDGVHNSNTQYIQISPTIIGTGDICISIVSGQ 280  
Db 234 QIHVSFDCNVVQASNLMTAPENSPTDGIHVTGQNIHSSCVIGTGDICISIVSGQ 293  
Qy 281 NVOATNTGCPGHGISIGSGNSSEAYVSNVTVNEAKIIGAENGVRKIKTWGGSGQASN 340  
Db 294 KAVVNDITGCPGHGISIGSGNSSEAYVSNVTVNEAKIIGAENGVRKIKTWGGSGQASN 353  
Qy 341 IKFLANEMODVKYPIIIDQNYCDRVEPCIQFSAVQVKNVYENIKGTSATKYAIKEDCS 400  
Db 354 IKFQNVEMHNVENPIIIDQNYCDQDKPCQEGSSAVQVKNVYQVNIKGTCA SNVAITFDGS 413  
Qy 401 TNFPCGIIEMNTLVGSGKPEATCKVNFNNAEHTPHCTSLSEDE 452  
Db 414 KRFPCGIIVEDVLEIEGGAALKCNVELSETGVSPHCEEGEE 462

## RESULT 4

Qy 094B15 PRELIMINARY; PRT; 444 AA.  
ID Q94B15  
AC Q94B15  
DT 01-DEC-2001 (TEMBLrel. 19, Created)  
DT 01-DEC-2001 (TEMBLrel. 19, Last sequence update)  
DT 01-MAR-2004 (TEMBLrel. 26, Last annotation update)  
DE Polyalacturonase PGL  
OS Vitis vinifera (Grape).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
OC Vitaceae; Vitis.  
OX NCBI\_TaxId=29760;  
RN (1)  
RP SEQUENCE FROM N.A.  
RC MEDLINE=21658619; PubMed=11800390;  
RA Nunan K.J., Davies C., Robinson S.P., Fincher G.B.;  
RT "Expression patterns of cell wall-modifying enzymes during grape berry  
RT development";  
RL Plant 214:257-264(2001).  
CC -1- SIMILARITY: Belongs to family 28 of glycosyl hydrolases.  
DR EMBL: AY043233; AAK81876.1; -;  
DR GO: GO:000518; C:cell wall; IEA.  
DR GO: GO:0016798; F:hydrolase activity, acting on glycosyl bonds; IEA.  
DR GO: GO:0004650; F:polygalacturonase activity; IEA.  
DR GO: GO:0005975; P:carbohydrate metabolism; IEA.  
DR InterPro: IPR000743; Glyco\_hydro\_28.  
DR InterPro: IPR006626; Pbh1.  
DR InterPro: IPR011050; Pectin\_lyase\_like.  
DR InterPro: IPR000408; Reg\_chit\_condens.  
DR Pfam: PF00295; Glyco\_hydro\_28; 1.  
DR SMART: SM00710; Pbh1; 4.  
DR PROSITE: PS00502; POLYGALACTURONASE; 1.  
DR PROSITE: PS00626; RCCL\_2; UNKNOWN\_1.  
DR Cell wall; Glycosidase; Hydrolase.  
KW SEQUENCE 444 AA; 47419 MW; D94259F47992F132 CRC64;

Query Match 50.4%; Score 1204.5; DB 2; Length 444;  
Best Local Similarity 61.5%; Pred. No. 4.5e-76;  
Matches 228; Conservative 53; Mismatches 87; Indels 3; Gaps 3;

Qy 73 IKVINLVSGANGDGKTYNIAEQAMNACSSRTPVQFVYVKNKYLKQITFGSPCRS 132  
Db 75 VKVNVNYYGAKGKDGSDATE-AFKAKMAACSSPGSV-LVVPKNKYLKQITFGSPCKS 132  
Qy 133 SISVKIFGLSEASKSISDY-KDRRLMIAPDSVQNLVVGCGGTNGGVWMPSSCKINKS 191  
Db 133 SITVQIVGVQASTDSASVSNDBTHTWLIIFENVNLAIVGGGTNGGKGTWESSCKVNYD 192  
Qy 192 LPCRDAPTALTFWNCIGLVNNLKSQAOIHKFESCTNVVNASNLMTINASAKSPNTDGV 251



```

Db 176 WPKSCKNPQLPCIGAPTAATVPECNRLRVSNIRLENAQOMHLTFQCKKVKALNLTWTS 235
Qy 242 SAKSPNTDGHVANTQYIOISDITIGTDCDICSIVSSQNVQATNTICGHHGISIGSLG 301
Db 236 PADSPTDGHVANTQYIOISDITIGTDCDICSIVSSQNVQATNTICGHHGISIGSLG 295
Qy 302 SGSEAVSVNTVNEAKIIGAENGVRKIKTQGGSGQASNIKPLNVEMQDYKYPPIIDONY 361
Db 296 EDNSEAVSVNTVNEAKIIGAENGVRKIKTQGGSGQASNIKPLNVEMQDYKYPPIIDONY 355
Qy 362 CDRVEPCIQQFSAVOQKNVYENIKGTSATKVAIKPDCSTNPFCEGIMENINLVESGK 421
Db 356 CDRVAPCEPKSAVOQKNVYENIKGTSATKVAIKPDCSTNPFCEGIMENINLVESGK 415
Qy 422 P-SEATCKNVHFNNAEHVPHCT 443
Db 416 DVSKASCSNVKLDTRGVNSPLCT 438

```

## RESULT 7

```

ID 08RY29 PRELIMINARY; PRT; 433 AA.
AC 08RY29;
DT 01-JUN-2002 (TEMBLrel. 21, Created)
DT 01-JUN-2002 (TEMBLrel. 21, Last sequence update)
DT 01-MAR-2004 (TEMBLrel. 26, Last annotation update)
DE AC2941850/11147.5.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosid II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Shim P., Chen H., Cheuk R., Kim C.-J., Meyers M.C., Banb U.,
RA Bower L., Carninci P., Chang E., Dale J.M., Goldsmith A.D.,
RA Hayashizaki Y., Ishida J., Jones T., Kamiya A., Karlin-Neumann G.,
RA Kawai J., Lam B., Lee J.M., Lin J., Miranda M., Narusaka M.,
RA Nguyen M., Onodera C.S., Palm C.J., Quach H.L., Sakurai T., Satou M.,
RA Seki M., Southwick A., Tang C.C., Toriumi M., Wu H.C., Yamada K.,
RA Yamamura Y., Yu G., Yu S., Shinozaki K., Davis R.W., Theologis A.,
RA Becker J.R.;
RL Submitted (FEF-2002) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Belongs to family 28 of glycosyl hydrolases.
CC EMBL: AY078936; AL84942.1; -.
DR GO: GO:0005618; C:cell wall; IEA.
DR GO: GO:0016798; F:hydrolase activity, acting on glycosyl bonds; IEA.
DR GO: GO:0004650; F:polysaccharonase activity; IEA.
DR GO: GO:0005975; P:carbohydrate metabolism; IEA.
DR InterPro: IPR000743; Glyco_hydro_28.
DR InterPro: IPR006626; Pbh1.
DR InterPro: IPR011050; Pectin_lyase_1like.
DR InterPro: IPR000408; Reg_chf_condense.
DR Pfam: PF00295; Glyco_hydro_28; 1.
DR SMART: SM00710; Pbh1_5.
DR PROSITE: PS00502; POLYGALACTURONASE; 1.
DR PROSITE: PS00628; RCT1_2; UNKNOWN_1.
DR Cell wall; Glycosidase; Hydrolase.
SQ SEQUENCE 433 AA; 46492 MW; 08E1EE62AA05A26B CRC64;

```

Query Match 48 1%; Score 1149; DB 2; Length 433;

Best Local Similarity 50.7%; Pred. No. 3.3e-72;

Matches 226; Conservative 76; Mismatches 118; Indels 26; Gaps 6;

```

Qy 6 NSILLIIIPASSISTCRSNVINDNLKFOY-----DNILEOFAPHPQAYLSYLSKNI 60
Db 6 NLTVFLWMLMLSWCKASRISPNVDHGYKPKSKSLKRR--EITGGRSPVRSRLR 63
Qy 61 SNNNDIKVNDKGIKIVNVLSFGAKGDKTYDNTAFEDANAEAGSSRPVQPVVPKKNYL 120
Db 64 TPT-----TVSVSDFGAKGDKTDDTQAFVNAWKKAGSSNGAVNLVLPKNTYL 112

```

```

Qy 121 LKQITFSGPCRSISYKIFGSLSEASGISDYKDRRLMIAFDVQNLVWGG--GTINGNG 178
Db 113 LKSIQLTGFCNSILTVQIFGTLSASQKRDYKDISKIMFDPGVNLSVDGGDGVVDGNG 172
Qy 179 QVWPPSSCKINSLPRDAPALTATFWCKRLKNNLSKSAQOIHIKFEECTVNASNL 238
Db 173 EFWMONSCRNKAKPCTXAPALTATFYNKSLYKNNKVRNAQOIISTECSNVQSVNV 232
Qy 239 INSAKSPNTDGHVANTQYIOISDITIGTDCDICSIVSSQNVQATNTICGHHGISIG 298
Db 233 VTAADSPNTDGHVANTQYIOISDITIGTDCDICSIVSSQNVQATNTICGHHGISIG 292
Qy 299 SLGSGNSEAVSVNTVNEAKIIGAENGVRKIKTQGGSGQASNIKPLNVEMQDYKYPPIID 358
Db 293 SLGDDNSKAFVSGVTVTDGAKLSTGDNVRIKTYQGGSGTASNIIFQNIQMDNVKXNPIID 352
Qy 359 QNYCDRVEPCIQQFSAVOQKNVYENIKGTSATKVAIKPDCSTNPFCEGIMENINLVGE 418
Db 353 QDDCDKQK-CTTEKSAVOQKNVYRDISGTSASENAITFNCSTKVPDQGIIVLDRVNIKG- 410
Qy 419 SGKPSSEATCKNVHFNNAEHVPHCTS 444
Db 411 ---GATCTNANVVDKGAVALPQCNS 432

```

## RESULT 8

PGAR\_PERAR STANDARD; PRT; 462 AA.

```

ID 002096;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Polyalacturonase precursor (EC 3.2.1.15) (PG) (Pectinase).
OS Persea americana (Avocado).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; magnoliids; Laurales; Lauraceae; Persea.
OX NCBI_TaxID=3435;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=cv. Hass; TISSUE=pericarp;
RX MEDLINE=93184201; PubMed=8095163;
RT Doiplo B., Lowe A.L., Wilson I.D., Merodio C., Giereson D.;
RT "Cloning and characterization of avocado fruit mRNAs and their
RT expression during ripening and low-temperature storage.";
RL Plant Mol. Biol. 21:437-449(1993).
RN [2]
RP SEQUENCE FROM N.A.
RA STRAIN=cv. Hass; TISSUE=Mesocarp; DOI=10.1104/pp.103.1.289;
RX MEDLINE=94269193; PubMed=8208850;
RA Kutsunai S.Y., Lin A.C., Percival F.W., Latties G.G.,
RA Christoffersen R.E.,
RT "Ripening-related polyalacturonase cDNA from avocado.";
RL Plant Physiol. 103:289-290(1993).
CC -1- FUNCTION: Acts in concert with the pectinesterase, in the ripening
CC process. Is involved in cell wall metabolism, specifically in
CC polyuronide degradation.
CC -1- CATALYTIC ACTIVITY: Random hydrolysis of 1,4-alpha-D-
CC galactosiduronic linkages in pectate and other galacturonans.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- DEVELOPMENTAL STAGE: In ripening fruit.
CC -1- SIMILARITY: Belongs to the glycosyl hydrolase 28 family.
CC -1- SIMILARITY: Contains 5 Pbh1 repeats.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.1sb-sib.ch/announce/
CC or send an email to license@1sb-sib.ch).
CC -----
DR EMBL: X66426; CAA47055.1; -.
DR EMBL: L06094; AAA32914.1; -.

```

DR EMBL: AJ250919; CAC05658.1; -.  
DR EMBL: AJ250918; CAC05657.1; -.  
DR GO: GO:0005618; C:cell wall; IEA.  
DR GO: GO:0016798; F:hydrolase activity, acting on glycosyl bonds; IEA.  
DR GO: GO:0004650; F:polygalacturonase activity; IEA.  
DR GO: GO:0005975; P:carbohydrate metabolism; IEA.  
DR InterPro: IPRO0743; Glyco\_hydro\_28.  
DR InterPro: IPRO06626; Pbh1.  
DR InterPro: IPRO11050; Pectin\_lyas\_like.  
DR InterPro: IPRO00408; Reg\_chf\_condens.  
DR Pfam: PF00295; Glyco\_hydro\_28; 1.  
DR SMART: SM00710; Pbh1\_5.  
DR PROSITE: PS00502; POLYGLACTURONASE; 1.  
DR PROSITE: PS00636; PCL\_2; UNKNOWN\_1.  
DR Cell wall; Glycosidase; Hydrolase.  
KW SEQUENCE 434 AA; 46603 MW; E9249AC1BEC0C219 CRC64;

| Query Match           | 47.3%            | Score 1130;   | DB 2;      | Length 434; |
|-----------------------|------------------|---|------------|-------------|
| Best Local Similarity | 49.8%            | Pred. No.7,1e-71;   |            |             |
| Matches 221;          | Conservative 78; | Mismatches 117;   | Indels 28; | Gaps 77;    |
| Qy                    | 8                | ILLIIITFPASSISTCTCSNVIDNLFKQYVDNIIIEQFPHADFCAYLS-VLSKRIEASNND   | 66         |             |
| Db                    | 11               | LLWSFLFMJSMCEASHNNI-----NRHYSVGTFSNSLIKRDDITRLK                 | 55         |             |
| Qy                    | 67               | KYDKNIRK---INVLSFGAKGDKTYNIIAFEOAMNACSSRTFVQFVBNKNRYLYLKQ       | 123        |             |
| Db                    | 56               | SVARASLRPTTVASVSDFGAKGDKGTDITQAFVANMKKACSSGAGVNIILVBEKTYFLKS    | 115        |             |
| Qy                    | 124              | IFSGPRCSBSIVKIFGSLFSAASKISDYDKRRLIIAFDSYONLVVGG--GTINGG--QV     | 180        |             |
| Db                    | 116              | ELNHPGCKSVTLVOLTGLTSLASQORSDYEDLSKITTFDGVNSLVLDGAGCTVANGNAET    | 175        |             |
| Qy                    | 181              | WWPSSCKINKSLPCRDAPALTALFPNCKNLKNNLKSRAAOQIIHKFESCTINVAASNLIN    | 240        |             |
| Db                    | 176              | WMENSCRENAKCKTAPALTALFINSKULKNNLRVADAQOIOISIEKCSNVVSNVEVT       | 235        |             |
| Qy                    | 241              | ASASPNPDGCHVSNTOYIOISDTIIIGTGDCLSIYSGSONVATNITCPGHSIGSL         | 300        |             |
| Db                    | 236              | APADSPNDGHIHTNTNIOVNSIIIGTGDDCLSISSGSONVINDLTCGPHGISISGL        | 295        |             |
| Qy                    | 301              | GSNGSEAYSVNVTYNEAKIIGAENGRIKTMOGSGSOASNIKFLANEMODVKYPIIIDON     | 360        |             |
| Db                    | 296              | GDDNSKAFVAGVTVDAKLSGTDNGRIRKTNQGGSGGTASNIIIFONTQMENVNPIIIDP     | 355        |             |
| Qy                    | 361              | YCPREVPCIOQFAVQKRVYVNIIGTSAIKVALKFDCLSTNPEEGILMININLVGSG        | 420        |             |
| Db                    | 356              | YCKK-SKCTEQKSAVQIKVYVRNIIISGTASADIALTFNCSKNYPCOGVILDKRVNIIKG--- | 411        |             |

|           |  |              |
|-----------|--|--------------|
| RESULT 10 |  |              |
| OB1246    |  |              |
| ID        | OB1246   |              |
| AC        | OB1246   | PRELIMINARY; |
| DT        |  | PRT; 461 AA. |
| DT        | 01-NOV-1998 (TREMBLrel. 08, Created)                                 |              |
| DT        | 01-NOV-1998 (TREMBLrel. 08, Last sequence update)                    |              |
| DT        | 01-MAR-2004 (TREMBLrel. 26, Last annotation update)                  |              |
| DE        | Polygalacturonase precursor.   |              |
| GN        | Name=MPG3;   |              |
| OS        | Cucumis melo (Muskmelon) .   |              |
| OC        | Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;   |              |
| OC        | Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; |              |
| OC        | eurosids I; Cucurbitales; Cucurbitaceae; Cucumis.                    |              |
| OX        | NCBI_TaxID=35656;  |              |
| RN        | [1]  |              |
| RP        | SEQUENCE FROM N.A.   |              |
| RX        | MEDLINE=98289082; PubMed=9635669; DOI=10.1104/jp.117.2.363;          |              |
| RA        | Hadfield K.A., Rose J.K., Yaver D.S., Berka R.M., Bennett A.B.:      |              |
| RT        | "Polygalacturonase precursor from Cucumis melo L. cv. 'Musk'."       |              |



RT for polygalacturonase in ripening-associated pectin disassembly.";  
 RL Plant Physiol. 117:363-373(1998).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Hadfield K.H., Rose J.K.C., Bennett A.B.;  
 RL Submitted (May-1998) to the EMBL/GenBank/DBJ databases.  
 CC - SIMILARITY: Belongs to family 28 of glycosyl hydrolases.  
 DR EMBL; AF062467; AAC26512.1; -  
 DR PIR; T08215; T08215.  
 DR GO; GO:0005618; C:cell wall; IEA.  
 DR GO; GO:0016798; F:hydrolase activity, acting on glycosyl bonds; IEA.  
 DR GO; GO:0004650; F:polygalacturonase activity; IEA.  
 DR GO; GO:0005975; P:carbohydrate metabolism; IEA.  
 DR InterPro; IPR00743; Glyco\_hydro\_28.  
 DR InterPro; IPR006626; Pbh1.  
 DR InterPro; IPR011050; Pectin\_lyase\_like.  
 DR InterPro; IPR00408; Reg\_ch1\_condens.  
 DR Pfam; PF00295; Glyco\_hydro\_28; 1.  
 DR SMART; SM00710; Pbh1; 4.  
 DR PROSITE; PS00502; POLYGALACTURONASE; 1.  
 DR PROSITE; PS00626; RCCL\_2; UNKNOWN\_1.  
 DR Cell wall; Glycosidase; Hydrolase; Signal.  
 FT SIGNAL 1 34 Potential.  
 FT CHAIN 35 461 polygalacturonase.  
 FT SEQUENCE 461 AA; 50697 MW; 358DBCB037C022B CRC64;  
 Query Match 47.2%; Score 1128; DB 2; Length 461;  
 Best Local Similarity 50.2%; Pred. No. 1,1e-70;  
 Matches 224; Conservative 79; Mismatches 129; Indels 14; Gaps 7;  
 QY 7 STILLIIIFASSISTCRSNVID-DNFLKQVYDNLLEQFAHDFQAYLSKNIESNNMI 65  
 DB 17 TFLFVVVNVNEDTSTGCGSPVDNPLPSISSGDKEXYGHYHSSSESSMLRT--RL 74  
 QY 66 DKVDKNGI--KVINLISFGAKDGGTYNDIAFEQANNEACSSRTPEQVYVKKXNYLTK 122  
 DB 75 EKVVSPLASPEIFPNVDYDGMGDGED-DTEAFKETWKDACSSTNAI-FLVPCDRVYHLK 132  
 QY 123 QIFPSGPRSSISVKTFGSLASKSISDY-KDRRLWIAFSPVONLVVGGGTTNGNQVM 161  
 DB 133 PTFSPGSPPLFLQLEGTYKASPHISDYEDKRRHWIIFONLLGLRVEGKIINGGRKW 192  
 QY 182 WPSCKINKSLPCRDPTALTLPNNCKLKYNNLKSNAQOIHIFESCITNVASNLMTNA 241  
 DB 193 KLNCKVKNKLPCKEAPLATVTCYCNLRKRGRLFRNAQOMLSFQCNVYKALNMTYA 252  
 QY 242 SAKSPNTDGVHVSNTQYIQTSDTIIIGTDGDISIVSGSNVQATNITGFGHSISIGSLG 301  
 DB 253 PGNSPMTDGIHVGTGTFIVKNCILMTGDICISIVSGSKVRAKGIICGPGHSISIGSLG 312  
 QY 302 SGNSEAVSVNVTYNEAKTIIGANGVRKTKWOGSGQASNKKFLNVEMQDYKPIIIDONY 361  
 DB 313 AGSSEAVSNVNVDTAKFSGTSGNVRKTKWOGSGQAYQNIIFONIVADNTNPIIINONY 372  
 QY 362 CDREVERCIQOFSAYOVQNVYENIKGTSATKVAIKFDCSTNPFCEGIMENINLV----- 416  
 DB 373 CDQKECTQADAVANVSNVYKNIKRGSSABAVKPCSGSVCCGILLDINLVHKGN 432  
 QY 417 GSGSKPSEATCKXVHFENAEHVTPHC 442  
 DB 433 DDKSAQAEASCKVKKMNRGRVSPQC 458  
 RESULT 11  
 ID 023147 PRELIMINARY; PRT; 431 AA.  
 AC 023147;  
 DT 01-JAN-1998 (TEMBLrel. 05, Last Created)  
 DT 01-JAN-1998 (TEMBLrel. 05, Last sequence update)  
 DT 05-JUL-2004 (TEMBLrel. 27, Last annotation update)  
 DE Endo-polygalacturonase (Ac3G57510).  
 GN Name=ADPgl; Synonym=Ac3G57510/T8H10\_110, T8H10\_110;  
 OS Arabidopsis thaliana (Mouse-ear cress).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;  
 OC eurosid II; Brassicales; Brassicaceae; Arabidopsis.  
 OX NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21377431; PubMed=11485203; DOI=10.1023/A:1010619002833;  
 RA Sander L., Child R., Ulvakov P., Albrechtsen M., Borckhardt B.;  
 RT "Analysis of a dehiscence zone endo-polygalacturonase in oilseed rape  
 RT (Brassica napus) and Arabidopsis thaliana: evidence for roles in cell  
 RT separation in dehiscence and abscission zones, and in stelar tissues  
 RT during pollen tube growth.";  
 RL Plant Mol. Biol. 46:469-479(2001).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Benes V., Rechmann S., Borkova D., Ansgore W., Mewes H.W., Lemcke K.,  
 RA Meyer K.F.X., Queciter F., Salanoubat M.;  
 RL Submitted (Nov-1999) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RA BU Arabidopsis sequencing project;  
 RL Submitted (Jan-2000) to the EMBL/GenBank/DBJ databases.  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RA Jenkins E.S., Paul W., Craze M., Whitelaw C.A., Weigand A.,  
 RA Roberts J.A.;  
 RT "Dehiscence-related expression of an Arabidopsis thaliana gene  
 RT encoding a polygalacturonase in transgenic plants of Brassica napus.";  
 RL Plant Cell Environ. 22:159-167(1999).  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RA Seki M., Iida K., Satou M., Sakurai T., Akiyama K., Ishida J.,  
 RA Nakajima M., Endo A., Kamiya A., Narusaka M., Carninci P., Kawai J.,  
 RA Hayashizaki Y., Shinozaki K.;  
 RL Submitted (Nov-2002) to the EMBL/GenBank/DBJ databases.  
 RN [6]  
 RP SEQUENCE FROM N.A.  
 RA Shim P., Chen H., Cheuk R., Kim C.J., Bowser L., Carninci P.,  
 RA Chan M.W., Chang C.H., Dale J.M., Hayashizaki Y., Huan V.W.,  
 RA Ishida J., Jones T., Kamiya A., Karlin-Neumann G., Kawai J., Lam B.,  
 RA Lee J.M., Lin J., Miranda M., Narusaka M., Nguyen M., Onodera C.S.,  
 RA Palm C.J., Quach H.L., Sakurai T., Satou M., Seki M., Southwick A.,  
 RA Tang C.C., Toriumi M., Wong C., Wu H.C., Yamada K., Yu G., Yuan S.,  
 RA Shinozaki K., Davis R.W., Theologis A., Becker J.R.;  
 RL Submitted (Mar-2003) to the EMBL/GenBank/DBJ databases.  
 CC - SIMILARITY: Belongs to family 28 of glycosyl hydrolases.  
 DR EMBL; AU002532; CA05525.1; -  
 DR EMBL; AL13248; CAB66108.1; -  
 DR EMBL; AF037367; AAC98923.1; -  
 DR EMBL; AK117942; BAC42580.1; -  
 DR EMBL; E005376; AAO63440.1; -  
 DR PIR; T46187; T46187.  
 DR GO; GO:0005618; C:cell wall; IEA.  
 DR GO; GO:0016798; F:hydrolase activity, acting on glycosyl bonds; IEA.  
 DR GO; GO:0004650; F:polygalacturonase activity; IEA.  
 DR GO; GO:0005975; P:carbohydrate metabolism; IEA.  
 DR InterPro; IPR00743; Glyco\_hydro\_28.  
 DR InterPro; IPR006626; Pbh1.  
 DR InterPro; IPR011050; Pectin\_lyase\_like.  
 DR InterPro; IPR00408; Reg\_ch1\_condens.  
 DR Pfam; PF00295; Glyco\_hydro\_28; 1.  
 DR SMART; SM00710; Pbh1; 5.  
 DR PROSITE; PS00502; POLYGALACTURONASE; 1.  
 DR PROSITE; PS00626; RCCL\_2; UNKNOWN\_1.  
 DR Cell wall; Glycosidase; Hydrolase.  
 FT SEQUENCE 431 AA; 46572 MW; 0F0F30DF45804FE7 CRC64;  
 Query Match 47.2%; Score 1127; DB 2; Length 431;  
 Best Local Similarity 49.6%; Pred. No. 1,1e-70;  
 Matches 223; Conservative 71; Mismatches 116; Indels 40; Gaps 6;  
 QY 7 STILLIIIFASSISTCRSNVID-DNFLKQVYDNLLEQFAHDFQAYLSYL 55  
 DB 17 TFLFVVVNVNEDTSTGCGSPVDNPLPSISSGDKEXYGHYHSSSESSMLRT--RL 74

```

Db          9  AVFLCVLMLSLCKALSSNVDDGVEHDEGSFESDSLKLKNDVLT-----SLISSD 59
Qy          56  SKNESNNNDKXDKNGIKYINVLSTFGAKGDKTYDNIAFEQANNEACSSRTPVQFVYVK 115
Db          60  ETLLEAS-----TVSVSNFGAKDGKTDQTOAFKAMKKACSTNGVTFLEVPK 107
Qy          116  NKNYLLKQITFGSPCRSSISVKIFGSLFASSKIIPYKDRRLMIAFDSVQNLVGGG--GT 173
Db          108  GKTYLLKSTRFRPCSLNPLQLGLSLSTRSKYKDKNMHLLLEDNNLSIDGSGSTGI 167
Qy          174  INNGGVWMPSSCKINKSLPCBDAPLALTFFMNCNKLKVNNLKSRNAQOIHKFESCTNVV 233
Db          168  INNGKTMQNSCKIDKSPCKTAPLALTLYNLKVLNKNLKVNAQOIISTEKKNVE 227
Qy          234  ASNLMLNSAKSPNTDGVHVSNTQYIOISDTIIGDDCISIVSGSNVQATNITGPGHG 293
Db          228  VNVEITVAGDSEPTDGIHTNTQINRSNDITGDDCISIEGTQNIQIFPLTCGPGH 287
Qy          294  GISIGLSGNSSEAVSVTVNEAKIIGAENGVRIKTMGGSGQASNIKFLANVEMQVYK 353
Db          288  GISIGSLDDNSKAVSGINVDGAKFESDNGVRIKTYGGSGTAKNIFQNIHRENVKN 347
Qy          354  PIIDONVCDRVEPCIOQPSAVQKVVYENIKGTSATKVAIKFDCSTNPFCEGIMENI 413
Db          348  PIIDODYCDK-DKCEQESAVQKVVYENIKGTSATDVALTLNCSERYPCOGIVLENV 406
Qy          414  NLVSGSGPSEATCKNVHFNNAEHVTPHCT 443
Db          407  KIKG-----GTASCKNANVNGVSPKCS 431

RESULT 12
ID 022935 PRELIMINARY; RT; 426 AA.
AC 022935;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, last annotation update)
DE Putative polygalacturonase.
GN Name=At2g41850;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Rounsley S.D., Ian X., Ketchum K.A., Crosby M.L., Brandon R.C.,
RA Spriggs T.A., Mason T.M., Kerlavage A.R., Adams M.D., Somerville C.R.,
RA Venter J.C.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Town C.D., Kaul S.;
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Belongs to the family 28 of glycosyl hydrolases.
DR EMBL; AC002339; AAC02763.1; -.
DR PIR; H84846; H84846.
DR GO; GO:0005618; C:cell wall; IEA.
DR GO; GO:0016798; F:hydrolase activity, acting on glycosyl bonds; IEA.
DR GO; GO:0004650; F:polygalacturonase activity; IEA.
DR GO; GO:0005975; F:carbohydrate metabolism; IEA.
DR InterPro; IPR000743; Glyco_hydro_28.
DR InterPro; IPR006626; pBH1.
DR InterPro; IPR011050; Pectin_lyas_like.
DR InterPro; IPR000408; Reg chr condens.
DR Pfam; PF00295; Glyco_hydro_28; 1.
DR SMART; SM00502; POLYGALACTURONASE; 1.
DR PROSITE; PS00626; RCCL_2; UNKOWN 1.
DR Cell wall; Glycosidase; Hydrolase-.
KM SEQUENCE 426 AA; 45924 MW; F756754C0390A80B CRC64;

```

```

Query Match 46.7%; Score 1115.5; DB 2; Length 426;
Best Local Similarity 49.8%; Pred. No. 7.2e-70;
Matches 222; Conservative 74; Mismatches 117; Gaps 7;

Qy          6  NSILLIIIPASSISTGCSNVINDNLEFKQY-----DNILEQFAFDFOAYLSYSLKNE 60
Db          6  NLVTFELMLALMPSWCKASRISPNVDHSYKRFKSDSLIKR--EDITGLASFPAIRAR 63
Qy          61  SNNIDKVDKNGIKYINVLSTFGAKGDKTYDNIAFEQANNEACSSRTPVQFVYVKRYL 120
Db          64  TPT-----TVSDPFGAKDGKTDQTOAFVNAWKACSSNGAVNLLVPGNTYL 112
Qy          121  LKQITFGSPCRSSISVKIFGSLFASKISDYDRRLMIAFDSVQNLVGGG--GTINGNG 178
Db          113  LKSIQLTLPQNSILTVQIFGTLSASQKRSYDIOIKWIFDGVNLSLDGDDTGVVDNG 172
Qy          179  QVWMPSSCKINKSLPCBDAPLALTFFMNCNKLKVNNLKSRNAQOIHKFESCTNVVSNLM 238
Db          173  ETWQNSCKRNKA-----KALTFFNSKSLVKNLKVRNAQOIISTEKCSNVQVSNV 225
Qy          239  INASAKSPMTDGVHVSNTQYIOISDTIIGDDCISIVSGSNVQATNITGPGHGISTG 298
Db          226  VTPADSEPTDGIHTNTQINRSNDITGDDCISIEGQNVQINDITCGPHGISTG 285
Qy          299  SLGSGNSEAVSVTVNEAKIIGAENGVRIKTMGGSGQASNIKFLANVEMQVYKPIID 358
Db          286  SLGDDNSKAFVSGVTVDAKSLGTDNGVRIKTYGGSGTAKNIFQNIHRENVKNPIID 345
Qy          359  ONYCDRVEPCIOQPSAVQKVVYENIKGTSATKVAIKFDCSTNPFCEGIMENINLVE 418
Db          346  QDYCDK-SKCTTESAAVQKVVYENIKGTSASBNALTFTCSNGVCCGIVLDRVNIKG- 403
Qy          419  SKPSEATCKNVHFNNAEHVTPHCTS 444
Db          404  ----GKATCTNANVVDKGVLPQNS 425

RESULT 13
ID 042399 PRELIMINARY; RT; 433 AA.
AC 042399;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, last annotation update)
DE Endo-polygalacturonidase (Polygalacturonase).
OS Brassica napus (Rape).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC NCBI_TaxID=3708;
RN [1]
RP SEQUENCE FROM N.A.
RA Sander L., Bortemann J., Ulvskov P., Borkhardt B.;
RA Petersen M., Sander L., Child R., van Onckelen H., Ulvskov P.,
RA Borkhardt B.;
RL "Isolation and characterization of a pod dehiscence zone specific
RT polygalacturonase from Brassica napus.";
RT Plant Mol. Biol. 31:517-527(1996).
CC -1- SIMILARITY: Belongs to family 28 of glycosyl hydrolases.
DR EMBL; X98373; CA67020.1; -.
DR EMBL; X95800; CA65072.1; -.
DR GO; GO:0005618; C:cell wall; IEA.
DR GO; GO:0016798; F:hydrolase activity, acting on glycosyl bonds; IEA.
DR GO; GO:0004650; F:polygalacturonase activity; IEA.
DR GO; GO:0005975; F:carbohydrate metabolism; IEA.
DR InterPro; IPR000743; Glyco_hydro_28.

```



DR InterPro: IPR006626; Pbh1.  
 DR InterPro: IPR011050; Pectin lyase like.  
 DR Pfam: PF00295; Glyco\_hydro\_28; 1.  
 DR SMART: SM00710; Pbh1; 6.  
 DR PROSITE: PS00502; POLYGALACTURONASE; 1.  
 DR PROSITE: PS00626; RCCL\_2; UNKNOWN\_1.  
 DR Cell wall; Glycosidase; Hydrolase.  
 KM SEQUENCE 433 AA; 46595 MW; 1EDB29ACB958260 CRC64;

Query Match 46.6%; Score 1113; DB 2; Length 433;

Best Local Similarity 49.9%; Pred. No. 1.1e-69;

Matches 223; Conservative 69; Mismatches 121; Indels 34; Gaps 8;

QY 7 SILLIIIPASSISTGRSNVIDDNLFRQVYDNLLEQFAHDPQAYLSYLSKNIENNNID 66  
 DB 9 AVELCVLLMLACQALSNV-DDG-----YGHEDG---SFERISLILKLNDD 51  
 QY 67 KV-----DKNGIKVINLVSGAGDQKTYDNIAPQAMNBAQSSRTPVQFVPEKNK 118  
 DB 52 DUTLTKSDRPTTESSTVSINFGAKGDKTDTQAFKAMKACSTNGVTTTLIRKGT 111  
 QY 119 YLLKQITFGSPCRSSISVKIFGSLBASKISDY-KDRLMIADSVQNLVVGSG--GTIN 175  
 DB 112 YLLKSIKRFQGPCSLMSFQILGTLASSTKSDYSNDKNHLLLEDVNNLSIDGSAIYD 171  
 QY 176 GNGQVMPSPSCIKNSLPCRDAPLTALTFNNCKLKYNNLKSQAQOIHIKFESECTNVAS 225  
 DB 172 GNGNIMWONSCKIDKSPCTKAPTALTLVNLKNLVNLRVNAOOIQISIEKCNVGVK 221  
 QY 236 NLMINASAPSPNDGVHVSNTQYIOISDTIIGTDDCISIVSGSNVQATINICGPHGI 295  
 DB 232 NKAITPGDSPTNDGHIHVAITKNIRISNSDIGTDCISIEDSONVQINDLTCGPHGI 291  
 QY 296 SIGSLGSGNSEAVSNVTYNEAKIIGAENGVRIKTWQSGSGQASNIKFLAVENQDVXPI 355  
 DB 292 SIGSLGDDNSKAVSGIDVDGATLSETDNGVRIKTWQSGSGTAKNIKFOIRMDVKNPI 351  
 QY 356 IIDQNVCDRVEPTIOQPSAVQVQNVYENIKGTSATKVAIKFPCSTNPFCEGIMENINL 415  
 DB 352 IIDQNVCDK-DKCEQGESAVQVQNVYQNIKGTSTADVALMFCNSVYKPCQGIIVLENVI 410  
 QY 416 VESGKPSSEATCKNHFNNAEHTPHC 442  
 DB 411 KG-----GRKSCENNVKDKGTVSPKC 432

## RESULT 14

Q43063 PRELIMINARY; PRT; 458 AA.

AC Q43063; TREMBLrel. 01, Created)  
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
 DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)  
 DE Polyalacturonase (EC 3.2.1.15).  
 OS Prunus persica (Peach).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
 OC eurosid I; Rosales; Rosaceae; Amygdaloideae; Prunus.  
 NCBI\_TaxID=3760;

RP SEQUENCE FROM N.A.  
 RA Lee E., Speirs J., Brady C.J.;  
 RT "Homologues to the tomato endopolgalacturonase gene in the peach  
 genome";  
 RL Plant Cell Environ. 13:513-521(1990).

RP MEDLINE=94302167; PubMed=8029352; DOI=10.1104/pp.105.1.125;  
 RA Leister D.R., Speiers J., Orr G., Brady C.J.;  
 RT "Peach (Prunus persica) endopolgalacturonase - cDNA isolation and mRNA  
 analysis in melting and nonmelting peach cultivars";  
 RL Plant Physiol. 105:125-231(1994).

RN [3]  
 RP SEQUENCE FROM N.A.

RA Speirs J.;

RU Submitted (APR-1994) to the EMBL/GenBank/DBJ databases.

CC -1- SIMILARITY: Belongs to family 28 of glycosyl hydrolases.

DR EMBL; X77231; CAA54448.1; -.

DR PIR; S71523; S71523.

DR GO; GO:0016798; F:hydrolyase activity, acting on glycosyl bonds; IEA.

DR GO; GO:0004650; F:polygalacturonase activity; IEA.

DR GO; GO:0005975; P:carbohydrate metabolism; IEA.

DR InterPro: IPR00743; Glyco\_hydro\_28.

DR InterPro: IPR006626; Pbh1.

DR InterPro: IPR011050; Pectin lyase like.

DR InterPro: IPR00408; Reg. chr. condens.

DR Pfam: PF00295; Glyco\_hydro\_28; 1.

DR SMART; SM00710; Pbh1; 6.

DR PROSITE; PS00502; POLYGALACTURONASE; 1.

DR PROSITE; PS00626; RCCL\_2; UNKNOWN\_1.

KM Cell wall; Glycosidase; Hydrolase.

SO SEQUENCE 458 AA; 49716 MW; E7BABECB8B68A5 CRC64;

Query Match 46.4%; Score 1108; DB 2; Length 458;  
 Best Local Similarity 48.9%; Pred. No. 2.7e-69;  
 Matches 226; Conservative 82; Mismatches 126; Indels 28; Gaps 8;

QY 1 MYIORSILLIIIPASSISTGRSNVIDDNLFRQVYDNLLEQFAHDPQAYLSYLSKNIENNNID 56  
 DB 1 MALQKHLVLFYVVSFCASCSYSGQEVNLSLSPDH--EKESGYNRAHPNNMTIEG 58  
 QY 57 -----KNIESNNIDKV---DKNGIKVINLVSGAGDQKTYDNIAPQAMNBAQSSRTPVQFVPEKNK 103  
 DB 59 VKMEFIKRAQLFSGKLEBAGSKSSSVKTSVANFGAKGAD-DTFAPEKAAAC 117  
 QY 104 SERTPVQFVPEKNKYLKQITFGSPCRSSISVKIFGSLBASKISDYKDRLMIADSV 163  
 DB 118 SSGALVAVLPQ-KTLVAPIESGPKSHLTQIIGTIEASDRSVYKDVTHLIFDNV 176  
 QY 164 ONLVVGGGTGNGQVMPSPSCIKNSLPCRD-APTALTFNNCKLKYNNLKSQAQOIHIKFESECTNVAS 222  
 DB 177 QSLVVGPGTNGNGRMHNSCKRFPQPCDQATATYFNKCNLVKYLQDAQGM 236  
 QY 177 QSLVVGPGTNGNGRMHNSCKRFPQPCDQATATYFNKCNLVKYLQDAQGM 236  
 QY 223 HIKFESECTNVASNLMINASAPSPNDGVHVSNTQYIOISDTIIGTDDCISIVSGSNV 282  
 DB 237 HVRFGCKVVEASHLVTAPEDSPTNDGHIHTVTKNITSSVIGTDDCISIVSGSNV 296  
 QY 283 QATNITCGPHGISIGSLGSGNSEAVSNVTYNEAKIIGAENGVRIKTWQSGSGQASNIK 342  
 DB 297 QATDITCGPHGISIGSLGEDNANDHVGVFVNGAKISGTSNGVRIKTWQSGSGASNTV 356  
 QY 343 FNVENQDVXPIIIDQNVCD-RVEPCIOQPSAVQVQNVYENIKGTSATKVAIKFPCST 401  
 DB 357 FQVENVNDVTNPIIIDQNVCDHKNKDCITRORSVQVKNVLYQNIKGTSTADITALTNCSQ 416  
 QY 402 NPFCEGIMENINLVGESGKPSSEATCKNHFNNAEHTPHC 443  
 DB 417 SVPCQGIIVQNIOL-----QNAKACNNVKNPVAKVASPRCS 453

## RESULT 15

Q8VXT3 PRELIMINARY; PRT; 433 AA.

AC Q8VXT3; TREMBLrel. 20, Created)

DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)

DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)

DE Endo polygalacturonase.

OS Brassica rapa (Turnip).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
 OC eurosid II; Brassicales; Brassicaceae; Brassica.  
 NCBI\_TaxID=51350;  
 RP SEQUENCE FROM N.A.



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 24, 2005, 06:59:05 ; Search time 484 Seconds

(without alignments)  
1102.849 Million cell updates/sec

Title: US-10-691-374-2

Perfect score: 2390  
Sequence: 1 MVIQRNILLIIFASIS.....VTPHCTSLISEDEALYNY 457

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 6959266 seqs, 116806243 residues

Total number of hits satisfying chosen parameters: 6959266

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 100%  
Listing first 45 summaries

Database : Pending Patents AA Main:\*

- 1: /cgn2\_6/ptodata/1/paa/PCTUS\_COMB.pep.\*
- 2: /cgn2\_6/ptodata/1/paa/US06\_COMB.pep.\*
- 3: /cgn2\_6/ptodata/1/paa/US08\_COMB.pep.\*
- 4: /cgn2\_6/ptodata/1/paa/US08\_COMB.pep.\*
- 5: /cgn2\_6/ptodata/1/paa/US08\_COMB.pep.\*
- 6: /cgn2\_6/ptodata/1/paa/US08\_COMB.pep.\*
- 7: /cgn2\_6/ptodata/1/paa/US08\_COMB.pep.\*
- 8: /cgn2\_6/ptodata/1/paa/US08\_COMB.pep.\*
- 9: /cgn2\_6/ptodata/1/paa/US08\_COMB.pep.\*
- 10: /cgn2\_6/ptodata/1/paa/US08\_COMB.pep.\*
- 11: /cgn2\_6/ptodata/1/paa/US08\_COMB.pep.\*
- 12: /cgn2\_6/ptodata/1/paa/US08\_COMB.pep.\*
- 13: /cgn2\_6/ptodata/1/paa/US08\_COMB.pep.\*
- 14: /cgn2\_6/ptodata/1/paa/US08\_COMB.pep.\*
- 15: /cgn2\_6/ptodata/1/paa/US08\_COMB.pep.\*
- 16: /cgn2\_6/ptodata/1/paa/US08\_COMB.pep.\*
- 17: /cgn2\_6/ptodata/1/paa/US08\_COMB.pep.\*
- 18: /cgn2\_6/ptodata/1/paa/US08\_COMB.pep.\*
- 19: /cgn2\_6/ptodata/1/paa/US08\_COMB.pep.\*
- 20: /cgn2\_6/ptodata/1/paa/US08\_COMB.pep.\*
- 21: /cgn2\_6/ptodata/1/paa/US08\_COMB.pep.\*
- 22: /cgn2\_6/ptodata/1/paa/US08\_COMB.pep.\*
- 23: /cgn2\_6/ptodata/1/paa/US08\_COMB.pep.\*
- 24: /cgn2\_6/ptodata/1/paa/US08\_COMB.pep.\*
- 25: /cgn2\_6/ptodata/1/paa/US08\_COMB.pep.\*
- 26: /cgn2\_6/ptodata/1/paa/US08\_COMB.pep.\*
- 27: /cgn2\_6/ptodata/1/paa/US08\_COMB.pep.\*
- 28: /cgn2\_6/ptodata/1/paa/US08\_COMB.pep.\*
- 29: /cgn2\_6/ptodata/1/paa/US08\_COMB.pep.\*
- 30: /cgn2\_6/ptodata/1/paa/US08\_COMB.pep.\*
- 31: /cgn2\_6/ptodata/1/paa/US08\_COMB.pep.\*
- 32: /cgn2\_6/ptodata/1/paa/US08\_COMB.pep.\*
- 33: /cgn2\_6/ptodata/1/paa/US08\_COMB.pep.\*
- 34: /cgn2\_6/ptodata/1/paa/US08\_COMB.pep.\*
- 35: /cgn2\_6/ptodata/1/paa/US08\_COMB.pep.\*
- 36: /cgn2\_6/ptodata/1/paa/US08\_COMB.pep.\*
- 37: /cgn2\_6/ptodata/1/paa/US08\_COMB.pep.\*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score  | Query Match | Length | DB ID                   | Description        |
|------------|--------|-------------|--------|-------------------------|--------------------|
| 1          | 2390   | 100.0       | 457    | 1 PCT-US03-37406-2      | Sequence 2, Appl1  |
| 2          | 2390   | 100.0       | 457    | 22 US-09-791-537-50958  | Sequence 50958, A  |
| 3          | 2390   | 100.0       | 457    | 26 US-10-018-604-2      | Sequence 2, Appl1  |
| 4          | 2390   | 100.0       | 457    | 32 US-10-691-374-2      | Sequence 2, Appl1  |
| 5          | 2382   | 99.7        | 457    | 1 PCT-US03-37406-48     | Sequence 48, Appl1 |
| 6          | 2382   | 99.7        | 457    | 1 PCT-US03-37406-50     | Sequence 50, Appl1 |
| 7          | 2382   | 99.7        | 457    | 32 US-10-691-374-48     | Sequence 48, Appl1 |
| 8          | 2382   | 99.7        | 457    | 32 US-10-691-374-50     | Sequence 50, Appl1 |
| 9          | 2338.5 | 97.8        | 456    | 1 PCT-US02-28066-1      | Sequence 1, Appl1  |
| 10         | 2338.5 | 97.8        | 456    | 22 US-09-791-537-102966 | Sequence 102966, A |
| 11         | 2338.5 | 97.8        | 456    | 22 US-10-487-753-1      | Sequence 1, Appl1  |
| 12         | 1295.5 | 54.2        | 467    | 22 US-09-791-537-49108  | Sequence 49108, A  |
| 13         | 1185   | 49.6        | 438    | 22 US-09-791-537-9525   | Sequence 9525, Ap  |
| 14         | 1153   | 48.2        | 453    | 22 US-09-791-537-83144  | Sequence 83144, A  |
| 15         | 1142   | 47.8        | 462    | 22 US-09-791-537-75907  | Sequence 75907, A  |
| 16         | 1136   | 47.5        | 462    | 22 US-09-791-537-85713  | Sequence 85713, A  |
| 17         | 1128   | 47.2        | 461    | 22 US-09-791-537-140414 | Sequence 140414, A |
| 18         | 1127   | 47.2        | 431    | 19 US-09-513-996A-51400 | Sequence 51400, A  |
| 19         | 1127   | 47.2        | 431    | 22 US-09-791-537-36586  | Sequence 36586, A  |
| 20         | 1127   | 47.2        | 431    | 33 US-10-739-930-6074   | Sequence 6074, Ap  |
| 21         | 1127   | 47.2        | 463    | 19 US-09-513-996A-51399 | Sequence 51399, A  |
| 22         | 1119.5 | 46.8        | 415    | 19 US-09-513-996A-51401 | Sequence 51401, A  |
| 23         | 1115.5 | 46.7        | 426    | 21 US-09-708-427-16483  | Sequence 16483, A  |
| 24         | 1115.5 | 46.7        | 426    | 22 US-09-791-537-7224   | Sequence 7224, Ap  |
| 25         | 1113.5 | 46.6        | 409    | 21 US-09-708-427-16484  | Sequence 16484, A  |
| 26         | 1113.5 | 46.6        | 433    | 22 US-09-791-537-115275 | Sequence 115275, A |
| 27         | 1108   | 46.4        | 458    | 22 US-09-791-537-17408  | Sequence 17408, A  |
| 28         | 1103   | 46.2        | 433    | 20 US-09-646-679-32     | Sequence 32, Appl  |
| 29         | 1103   | 46.2        | 433    | 22 US-09-791-537-101225 | Sequence 101225, A |
| 30         | 1103   | 46.2        | 433    | 33 US-10-739-930-7092   | Sequence 7092, Ap  |
| 31         | 1103   | 46.2        | 433    | 33 US-10-787-986-32     | Sequence 32, Appl  |
| 32         | 1091   | 45.6        | 460    | 22 US-09-791-537-52113  | Sequence 52113, A  |
| 33         | 1061.5 | 44.4        | 335    | 22 US-09-791-537-70434  | Sequence 70434, A  |
| 34         | 1058.5 | 44.3        | 404    | 22 US-09-791-537-17410  | Sequence 17410, A  |
| 35         | 1011   | 42.3        | 393    | 30 US-10-437-963-127090 | Sequence 127090, A |
| 36         | 1010   | 42.3        | 452    | 30 US-10-437-963-127091 | Sequence 127091, A |
| 37         | 988.5  | 41.4        | 444    | 30 US-10-437-963-111044 | Sequence 111044, A |
| 38         | 940    | 39.3        | 340    | 30 US-10-425-115-218415 | Sequence 218415, A |
| 39         | 939.5  | 39.3        | 433    | 37 US-60-324-109-20499  | Sequence 20499, A  |
| 40         | 925    | 38.7        | 452    | 1 PCT-US00-26884-22     | Sequence 22, Appl  |
| 41         | 925    | 38.7        | 452    | 18 US-09-410-191-22     | Sequence 22, Appl  |
| 42         | 925    | 38.7        | 452    | 22 US-09-791-537-54401  | Sequence 54401, A  |
| 43         | 911    | 38.1        | 438    | 30 US-10-449-907-36629  | Sequence 36629, A  |
| 44         | 898.5  | 37.6        | 459    | 22 US-09-791-537-123232 | Sequence 123232, A |
| 45         | 898.5  | 37.6        | 459    | 24 US-09-935-625-649    | Sequence 649, App  |

ALIGNMENTS

RESULT 1  
PCT-US03-37406-2

Sequence 2 Application PC/TUS0337406

GENERAL INFORMATION:

APPLICANT: McCallum, Claire

APPLICANT: Slade, Ann J.

APPLICANT: Colbert, Trent

APPLICANT: Knauf, Vic

TITLE OF INVENTION: Tomatoes Having Reduced Polygalacturonase Activity Caused by Non-

FILE REFERENCE: MBHB 02-276

CURRENT APPLICATION NUMBER: PCT/US03/37406

CURRENT FILING DATE: 2003-11-21

NUMBER OF SEQ ID NOS: 50

SOFTWARE: PatentIn version 3.1

SEQ ID NO 2

LENGTH: 457

TYPE: PRT

ORGANISM: Lycopersicon esculentum  
PCT-US03-37406-2

Query Match  
Best Local Similarity 100.0%; Score 2390; DB 1; Length 457;  
Matches 457; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 MVIQRNSILLIIIFASSISTCRSNVIDNLFKQYVDNILEQFAHDPQAVLSYLSKNIE 60
DB 1 MVIQRNSILLIIIFASSISTCRSNVIDNLFKQYVDNILEQFAHDPQAVLSYLSKNIE 60
QY 61 SNNNDKVDKNGIKYINVLSPGAKGDKTYDNIAFEQAMNEACSSRTPVQFVVPKKNY 120
DB 61 SNNNDKVDKNGIKYINVLSPGAKGDKTYDNIAFEQAMNEACSSRTPVQFVVPKKNY 120
QY 121 LKQITFSGPCRSSISVKIFGSLFASSTKIDYKDRRLMIADSVQNLVGGGGTINNGQV 180
DB 121 LKQITFSGPCRSSISVKIFGSLFASSTKIDYKDRRLMIADSVQNLVGGGGTINNGQV 180
QY 181 WMPSSCKINKSLPCRDAPALTFFWNCNKLKVNNLKSKNAQOIHIFKESCTNVVASNLMIN 240
DB 181 WMPSSCKINKSLPCRDAPALTFFWNCNKLKVNNLKSKNAQOIHIFKESCTNVVASNLMIN 240
QY 241 ASAKSPNTDGVHVSNTQYIQISDTIIGTDDCTSIYSGSQNVQATNITCGPHGISISGL 300
DB 241 ASAKSPNTDGVHVSNTQYIQISDTIIGTDDCTSIYSGSQNVQATNITCGPHGISISGL 300
QY 301 GSGNSEAVYNSVTNVEAKIIGAENGVRITKWQSGSQASNIKFLVEMQDVKYPPIIDON 360
DB 301 GSGNSEAVYNSVTNVEAKIIGAENGVRITKWQSGSQASNIKFLVEMQDVKYPPIIDON 360
QY 361 YCDRVEPCIOQFSAVQVKNVYENIKGTSATKVAIKFDCSTNFPCEGIMENINLVGESSG 420
DB 361 YCDRVEPCIOQFSAVQVKNVYENIKGTSATKVAIKFDCSTNFPCEGIMENINLVGESSG 420
QY 421 KPESEATCKNVHFNMAEHVTPHCTSLSEDEALLYNY 457
DB 421 KPESEATCKNVHFNMAEHVTPHCTSLSEDEALLYNY 457

```

## RESULT 2

US-09-791-537-50958  
Sequence 50958, Application US/09791537  
GENERAL INFORMATION:  
APPLICANT: Biomolix, Inc.  
APPLICANT: Dede, Derek  
TITLE OF INVENTION: Danzer, Joseph  
TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBE  
FILE REFERENCE: 261/210  
CURRENT APPLICATION NUMBER: US/09/791,537  
CURRENT FILING DATE: 2001-02-22  
NUMBER OF SEQ ID NOS: 153055  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 50958  
LENGTH: 457  
TYPE: PRT  
ORGANISM: Lycopersicon esculentum  
US-09-791-537-50958

Query Match  
Best Local Similarity 100.0%; Score 2390; DB 22; Length 457;  
Matches 457; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 MVIQRNSILLIIIFASSISTCRSNVIDNLFKQYVDNILEQFAHDPQAVLSYLSKNIE 60
DB 1 MVIQRNSILLIIIFASSISTCRSNVIDNLFKQYVDNILEQFAHDPQAVLSYLSKNIE 60
QY 61 SNNNDKVDKNGIKYINVLSPGAKGDKTYDNIAFEQAMNEACSSRTPVQFVVPKKNY 120
DB 61 SNNNDKVDKNGIKYINVLSPGAKGDKTYDNIAFEQAMNEACSSRTPVQFVVPKKNY 120
QY 121 LKQITFSGPCRSSISVKIFGSLFASSTKIDYKDRRLMIADSVQNLVGGGGTINNGQV 180
DB 121 LKQITFSGPCRSSISVKIFGSLFASSTKIDYKDRRLMIADSVQNLVGGGGTINNGQV 180

```

```

DB 121 LKQITFSGPCRSSISVKIFGSLFASSTKIDYKDRRLMIADSVQNLVGGGGTINNGQV 180
QY 181 WMPSSCKINKSLPCRDAPALTFFWNCNKLKVNNLKSKNAQOIHIFKESCTNVVASNLMIN 240
DB 181 WMPSSCKINKSLPCRDAPALTFFWNCNKLKVNNLKSKNAQOIHIFKESCTNVVASNLMIN 240
QY 241 ASAKSPNTDGVHVSNTQYIQISDTIIGTDDCTSIYSGSQNVQATNITCGPHGISISGL 300
DB 241 ASAKSPNTDGVHVSNTQYIQISDTIIGTDDCTSIYSGSQNVQATNITCGPHGISISGL 300
QY 301 GSGNSEAVYNSVTNVEAKIIGAENGVRITKWQSGSQASNIKFLVEMQDVKYPPIIDON 360
DB 301 GSGNSEAVYNSVTNVEAKIIGAENGVRITKWQSGSQASNIKFLVEMQDVKYPPIIDON 360
QY 361 YCDRVEPCIOQFSAVQVKNVYENIKGTSATKVAIKFDCSTNFPCEGIMENINLVGESSG 420
DB 361 YCDRVEPCIOQFSAVQVKNVYENIKGTSATKVAIKFDCSTNFPCEGIMENINLVGESSG 420
QY 421 KPESEATCKNVHFNMAEHVTPHCTSLSEDEALLYNY 457
DB 421 KPESEATCKNVHFNMAEHVTPHCTSLSEDEALLYNY 457

```

## RESULT 3

US-10-018-604-2  
Sequence 2, Application us/10018604  
GENERAL INFORMATION:  
APPLICANT: DANISCO A/S  
TITLE OF INVENTION: Process for the Enzymatic Modification of Pectin  
FILE REFERENCE: 550-354  
CURRENT APPLICATION NUMBER: US/10/018,604  
CURRENT FILING DATE: 2002-09-06  
NUMBER OF SEQ ID NOS: 8  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 2  
LENGTH: 457  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: PG enzyme  
US-10-018-604-2

Query Match  
Best Local Similarity 100.0%; Score 2390; DB 26; Length 457;  
Matches 457; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 MVIQRNSILLIIIFASSISTCRSNVIDNLFKQYVDNILEQFAHDPQAVLSYLSKNIE 60
DB 1 MVIQRNSILLIIIFASSISTCRSNVIDNLFKQYVDNILEQFAHDPQAVLSYLSKNIE 60
QY 61 SNNNDKVDKNGIKYINVLSPGAKGDKTYDNIAFEQAMNEACSSRTPVQFVVPKKNY 120
DB 61 SNNNDKVDKNGIKYINVLSPGAKGDKTYDNIAFEQAMNEACSSRTPVQFVVPKKNY 120
QY 121 LKQITFSGPCRSSISVKIFGSLFASSTKIDYKDRRLMIADSVQNLVGGGGTINNGQV 180
DB 121 LKQITFSGPCRSSISVKIFGSLFASSTKIDYKDRRLMIADSVQNLVGGGGTINNGQV 180
QY 181 WMPSSCKINKSLPCRDAPALTFFWNCNKLKVNNLKSKNAQOIHIFKESCTNVVASNLMIN 240
DB 181 WMPSSCKINKSLPCRDAPALTFFWNCNKLKVNNLKSKNAQOIHIFKESCTNVVASNLMIN 240
QY 241 ASAKSPNTDGVHVSNTQYIQISDTIIGTDDCTSIYSGSQNVQATNITCGPHGISISGL 300
DB 241 ASAKSPNTDGVHVSNTQYIQISDTIIGTDDCTSIYSGSQNVQATNITCGPHGISISGL 300
QY 301 GSGNSEAVYNSVTNVEAKIIGAENGVRITKWQSGSQASNIKFLVEMQDVKYPPIIDON 360
DB 301 GSGNSEAVYNSVTNVEAKIIGAENGVRITKWQSGSQASNIKFLVEMQDVKYPPIIDON 360
QY 361 YCDRVEPCIOQFSAVQVKNVYENIKGTSATKVAIKFDCSTNFPCEGIMENINLVGESSG 420
DB 361 YCDRVEPCIOQFSAVQVKNVYENIKGTSATKVAIKFDCSTNFPCEGIMENINLVGESSG 420

```

Db 361 YCDRVEPCIOQFSAVOVKNNVYENIKGTSATKVAIKFDCSTNPFCEGIIMENINLVGSEG 420  
Qy 421 KPSEATCKNVHFNNAEHVTPHCTSLSEI SEDEALLYNY 457  
Db 421 KPSEATCKNVHFNNAEHVTPHCTSLSEI SEDEALLYNY 457

RESULT 4  
US-10-691-374-2

Sequence 2. Application US/10691374  
GENERAL INFORMATION:  
APPLICANT: McCallum, Claire  
APPLICANT: Slade, Ann J.  
APPLICANT: Colbert, Trent  
APPLICANT: Knauf, Vic  
APPLICANT: Anaiah Inc.  
TITLE OF INVENTION: Tomatoes Having Reduced Polygalacturonase Activity Caused by Non-  
FILE REFERENCE: MBHB 02-276  
CURRENT APPLICATION NUMBER: US/10/691,374  
CURRENT FILING DATE: 2003-10-22  
NUMBER OF SEQ ID NOS: 50  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 2  
LENGTH: 457  
TYPE: PRT  
ORGANISM: Lycopersicon esculentum  
US-10-691-374-2

Query Match 100.0%; Score 2390; DB 32; Length 457;  
Best Local Similarity 100.0%; Pred. No. 1.2e-229;  
Matches 457; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MVTQRNSILLIIIFASSISTCRSNVIDNLFKQYVDNILEQEFADFOAYLSYSKNIE 60  
Db 1 MVTQRNSILLIIIFASSISTCRSNVIDNLFKQYVDNILEQEFADFOAYLSYSKNIE 60  
Qy 61 SNNNIDKVDKNGIKVINLVSFGAKGDKTYDNI AFQANNEACSSRTPVQFVVPKKNYL 120  
Db 61 SNNNIDKVDKNGIKVINLVSFGAKGDKTYDNI AFQANNEACSSRTPVQFVVPKKNYL 120  
Qy 121 LKQITSGPCRSISYKIFGSLASASKISDYKDRRLMIADSVQNLVVGCGGTINGNGQV 180  
Db 121 LKQITSGPCRSISYKIFGSLASASKISDYKDRRLMIADSVQNLVVGCGGTINGNGQV 180  
Qy 181 WMPSSCKINKSLPCRDAPFALTFMNCNKLKVNKLKSKNAQOIHIIKFESECTNVVASNLMIN 240  
Db 181 WMPSSCKINKSLPCRDAPFALTFMNCNKLKVNKLKSKNAQOIHIIKFESECTNVVASNLMIN 240  
Qy 241 ASAKSPNTDGVHVSNTQYIQISDTIIGTDDCISIVSGSQNVQATNITCGPHGISIGSL 300  
Db 241 ASAKSPNTDGVHVSNTQYIQISDTIIGTDDCISIVSGSQNVQATNITCGPHGISIGSL 300  
Qy 301 GSGNSEAYVSNVTVNEAKIIGAENGVRIKTWQSGSGQASNIKFLANVEMQVKKYPIIIDON 360  
Db 301 GSGNSEAYVSNVTVNEAKIIGAENGVRIKTWQSGSGQASNIKFLANVEMQVKKYPIIIDON 360  
Qy 361 YCDRVEPCIOQFSAVOVKNNVYENIKGTSATKVAIKFDCSTNPFCEGIIMENINLVGSEG 420  
Db 361 YCDRVEPCIOQFSAVOVKNNVYENIKGTSATKVAIKFDCSTNPFCEGIIMENINLVGSEG 420  
Qy 421 KPSEATCKNVHFNNAEHVTPHCTSLSEI SEDEALLYNY 457  
Db 421 KPSEATCKNVHFNNAEHVTPHCTSLSEI SEDEALLYNY 457

RESULT 5  
PCT-US03-37406-48  
Sequence 48. Application PC/TUS0337406  
GENERAL INFORMATION:  
APPLICANT: McCallum, Claire  
APPLICANT: Slade, Ann J.  
APPLICANT: Colbert, Trent

APPLICANT: Knauf, Vic  
APPLICANT: Anaiah Inc.  
TITLE OF INVENTION: Tomatoes Having Reduced Polygalacturonase Activity Caused by Non-  
FILE REFERENCE: MBHB 02-276  
CURRENT APPLICATION NUMBER: PCT/US03/37406  
CURRENT FILING DATE: 2003-11-21  
NUMBER OF SEQ ID NOS: 50  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 48  
LENGTH: 457  
TYPE: PRT  
ORGANISM: Lycopersicon esculentum  
PCT-US03-37406-48

Query Match 99.7%; Score 2382; DB 1; Length 457;  
Best Local Similarity 99.8%; Pred. No. 7.8e-229;  
Matches 456; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MVTQRNSILLIIIFASSISTCRSNVIDNLFKQYVDNILEQEFADFOAYLSYSKNIE 60  
Db 1 MVTQRNSILLIIIFASSISTCRSNVIDNLFKQYVDNILEQEFADFOAYLSYSKNIE 60  
Qy 61 SNNNIDKVDKNGIKVINLVSFGAKGDKTYDNI AFQANNEACSSRTPVQFVVPKKNYL 120  
Db 61 SNNNIDKVDKNGIKVINLVSFGAKGDKTYDNI AFQANNEACSSRTPVQFVVPKKNYL 120  
Qy 121 LKQITSGPCRSISYKIFGSLASASKISDYKDRRLMIADSVQNLVVGCGGTINGNGQV 180  
Db 121 LKQITSGPCRSISYKIFGSLASASKISDYKDRRLMIADSVQNLVVGCGGTINGNGQV 180  
Qy 181 WMPSSCKINKSLPCRDAPFALTFMNCNKLKVNKLKSKNAQOIHIIKFESECTNVVASNLMIN 240  
Db 181 WMPSSCKINKSLPCRDAPFALTFMNCNKLKVNKLKSKNAQOIHIIKFESECTNVVASNLMIN 240  
Qy 241 ASAKSPNTDGVHVSNTQYIQISDTIIGTDDCISIVSGSQNVQATNITCGPHGISIGSL 300  
Db 241 ASAKSPNTDGVHVSNTQYIQISDTIIGTDDCISIVSGSQNVQATNITCGPHGISIGSL 300  
Qy 301 GSGNSEAYVSNVTVNEAKIIGAENGVRIKTWQSGSGQASNIKFLANVEMQVKKYPIIIDON 360  
Db 301 GSGNSEAYVSNVTVNEAKIIGAENGVRIKTWQSGSGQASNIKFLANVEMQVKKYPIIIDON 360  
Qy 361 YCDRVEPCIOQFSAVOVKNNVYENIKGTSATKVAIKFDCSTNPFCEGIIMENINLVGSEG 420  
Db 361 YCDRVEPCIOQFSAVOVKNNVYENIKGTSATKVAIKFDCSTNPFCEGIIMENINLVGSEG 420  
Qy 421 KPSEATCKNVHFNNAEHVTPHCTSLSEI SEDEALLYNY 457  
Db 421 KPSEATCKNVHFNNAEHVTPHCTSLSEI SEDEALLYNY 457

RESULT 6  
PCT-US03-37406-50  
Sequence 50. Application PC/TUS0337406  
GENERAL INFORMATION:  
APPLICANT: McCallum, Claire  
APPLICANT: Slade, Ann J.  
APPLICANT: Colbert, Trent  
APPLICANT: Knauf, Vic  
APPLICANT: Anaiah Inc.  
TITLE OF INVENTION: Tomatoes Having Reduced Polygalacturonase Activity Caused by Non-  
FILE REFERENCE: MBHB 02-276  
CURRENT APPLICATION NUMBER: PCT/US03/37406  
CURRENT FILING DATE: 2003-11-21  
NUMBER OF SEQ ID NOS: 50  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 50  
LENGTH: 457  
TYPE: PRT  
ORGANISM: Lycopersicon esculentum  
PCT-US03-37406-50

Query Match 99.7%; Score 2382; DB 1; Length 457;  
 Best Local Similarity 99.8%; Pred. No. 7,8e-229;  
 Matches 456; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

QY 1 MVIORNSILLLIIIFASSISTCRSNVIDNLFKQVYDNILBOEFADHPQAYLSYSKNTIE 60
DB 1 MVIORNSILLLIIIFASSISTCRSNVIDNLFKQVYDNILBOEFADHPQAYLSYSKNTIE 60
QY 61 SNNNIDKVDKNGIKVINVLSPGAKGDKTYDNIAPFOANNEACSSRTVPQFVPPKKNKYL 120
DB 61 SNNNIDKVDKNGIKVINVLSPGAKGDKTYDNIAPFOANNEACSSRTVPQFVPPKKNKYL 120
QY 121 LKQITFSGPCRSSISVKIFGSLBASISKISDYKDRMLMIAFDSVQNLVVGGGGTINGNOV 180
DB 121 LKQITFSGPCRSSISVKIFGSLBASISKISDYKDRMLMIAFDSVQNLVVGGGGTINGNOV 180
QY 181 WMPSSCKINKSLPCRDAPALTFMNCNKLKVNNLKSKNAQOIHIFESCTNVVASNLMIN 240
DB 181 WMPSSCKINKSLPCRDAPALTFMNCNKLKVNNLKSKNAQOIHIFESCTNVVASNLMIN 240
QY 241 ASAKSPNTDGVHVSNTQYIISDPTIIGTGDDCISIVSGSQVQATNITGCGHGIGISGL 300
DB 241 ASAKSPNTDGVHVSNTQYIISDPTIIGTGDDCISIVSGSQVQATNITGCGHGIGISGL 300
QY 301 GSGNSEAVYNSVTVAEAKIIGAENGVRITKMGSSGQASNIKFLVEMQDVKYPPIIIDON 360
DB 301 GSGNSEAVYNSVTVAEAKIIGAENGVRITKMGSSGQASNIKFLVEMQDVKYPPIIIDON 360
QY 361 YCDRVEPCIQOFSAVQKVVYENIKGTSATKVAIKFPCSTNFPCEGIMENINLVGESSG 420
DB 361 YCDRVEPCIQOFSAVQKVVYENIKGTSATKVAIKFPCSTNFPCEGIMENINLVGESSG 420
QY 421 KPESEATCKNVHFNNAEHVTPHCTSLSEISDEALALYNY 457
DB 421 KPESEATCKNVHFNNAEHVTPHCTSLSEISDEALALYNY 457

RESULT 7
US-10-691-374-48
; Sequence 48, Application US/10691374
; GENERAL INFORMATION:
; APPLICANT: McCallum, Claire
; APPLICANT: Slade, Ann J.
; APPLICANT: Colbert, Trent
; APPLICANT: Knauf, Vic
; APPLICANT: Anawah Inc.
; TITLE OF INVENTION: Tomatoes Having Reduced Polygalacturonase Activity Caused by Non-
; FILE REFERENCE: MHBH 02-276
; CURRENT APPLICATION NUMBER: US/10/691,374
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 48
; LENGTH: 457
; TYPE: PRT
; ORGANISM: Lycopersicon esculentum
US-10-691-374-48

```

Query Match 99.7%; Score 2382; DB 32; Length 457;  
 Best Local Similarity 99.8%; Pred. No. 7,8e-229;  
 Matches 456; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

QY 1 MVIORNSILLLIIIFASSISTCRSNVIDNLFKQVYDNILBOEFADHPQAYLSYSKNTIE 60
DB 1 MVIORNSILLLIIIFASSISTCRSNVIDNLFKQVYDNILBOEFADHPQAYLSYSKNTIE 60
QY 61 SNNNIDKVDKNGIKVINVLSPGAKGDKTYDNIAPFOANNEACSSRTVPQFVPPKKNKYL 120
DB 61 SNNNIDKVDKNGIKVINVLSPGAKGDKTYDNIAPFOANNEACSSRTVPQFVPPKKNKYL 120
QY 121 LKQITFSGPCRSSISVKIFGSLBASISKISDYKDRMLMIAFDSVQNLVVGGGGTINGNOV 180
DB 121 LKQITFSGPCRSSISVKIFGSLBASISKISDYKDRMLMIAFDSVQNLVVGGGGTINGNOV 180

```

```

DB 121 LKQITFSGPCRSSISVKIFGSLBASISKISDYKDRMLMIAFDSVQNLVVGGGGTINGNOV 180
QY 181 WMPSSCKINKSLPCRDAPALTFMNCNKLKVNNLKSKNAQOIHIFESCTNVVASNLMIN 240
DB 181 WMPSSCKINKSLPCRDAPALTFMNCNKLKVNNLKSKNAQOIHIFESCTNVVASNLMIN 240
QY 241 ASAKSPNTDGVHVSNTQYIISDPTIIGTGDDCISIVSGSQVQATNITGCGHGIGISGL 300
DB 241 ASAKSPNTDGVHVSNTQYIISDPTIIGTGDDCISIVSGSQVQATNITGCGHGIGISGL 300
QY 301 GSGNSEAVYNSVTVAEAKIIGAENGVRITKMGSSGQASNIKFLVEMQDVKYPPIIIDON 360
DB 301 GSGNSEAVYNSVTVAEAKIIGAENGVRITKMGSSGQASNIKFLVEMQDVKYPPIIIDON 360
QY 361 YCDRVEPCIQOFSAVQKVVYENIKGTSATKVAIKFPCSTNFPCEGIMENINLVGESSG 420
DB 361 YCDRVEPCIQOFSAVQKVVYENIKGTSATKVAIKFPCSTNFPCEGIMENINLVGESSG 420
QY 421 KPESEATCKNVHFNNAEHVTPHCTSLSEISDEALALYNY 457
DB 421 KPESEATCKNVHFNNAEHVTPHCTSLSEISDEALALYNY 457

```

```

RESULT 8
US-10-691-374-50
; Sequence 50, Application US/10691374
; GENERAL INFORMATION:
; APPLICANT: McCallum, Claire
; APPLICANT: Slade, Ann J.
; APPLICANT: Colbert, Trent
; APPLICANT: Knauf, Vic
; APPLICANT: Anawah Inc.
; TITLE OF INVENTION: Tomatoes Having Reduced Polygalacturonase Activity Caused by Non-
; FILE REFERENCE: MHBH 02-276
; CURRENT APPLICATION NUMBER: US/10/691,374
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 50
; LENGTH: 457
; TYPE: PRT
; ORGANISM: Lycopersicon esculentum
US-10-691-374-50

```

Query Match 99.7%; Score 2382; DB 32; Length 457;  
 Best Local Similarity 99.8%; Pred. No. 7,8e-229;  
 Matches 456; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

QY 1 MVIORNSILLLIIIFASSISTCRSNVIDNLFKQVYDNILBOEFADHPQAYLSYSKNTIE 60
DB 1 MVIORNSILLLIIIFASSISTCRSNVIDNLFKQVYDNILBOEFADHPQAYLSYSKNTIE 60
QY 61 SNNNIDKVDKNGIKVINVLSPGAKGDKTYDNIAPFOANNEACSSRTVPQFVPPKKNKYL 120
DB 61 SNNNIDKVDKNGIKVINVLSPGAKGDKTYDNIAPFOANNEACSSRTVPQFVPPKKNKYL 120
QY 121 LKQITFSGPCRSSISVKIFGSLBASISKISDYKDRMLMIAFDSVQNLVVGGGGTINGNOV 180
DB 121 LKQITFSGPCRSSISVKIFGSLBASISKISDYKDRMLMIAFDSVQNLVVGGGGTINGNOV 180
QY 181 WMPSSCKINKSLPCRDAPALTFMNCNKLKVNNLKSKNAQOIHIFESCTNVVASNLMIN 240
DB 181 WMPSSCKINKSLPCRDAPALTFMNCNKLKVNNLKSKNAQOIHIFESCTNVVASNLMIN 240
QY 241 ASAKSPNTDGVHVSNTQYIISDPTIIGTGDDCISIVSGSQVQATNITGCGHGIGISGL 300
DB 241 ASAKSPNTDGVHVSNTQYIISDPTIIGTGDDCISIVSGSQVQATNITGCGHGIGISGL 300
QY 301 GSGNSEAVYNSVTVAEAKIIGAENGVRITKMGSSGQASNIKFLVEMQDVKYPPIIIDON 360
DB 301 GSGNSEAVYNSVTVAEAKIIGAENGVRITKMGSSGQASNIKFLVEMQDVKYPPIIIDON 360

```





Query Match 97.8%; Score 2338.5; DB 30; Length 456;  
Best Local Similarity 98.9%; Pred. No. 1.8e-224;  
Matches 452; Conservative 0; Mismatches 4; Indels 1; Gaps 1;

```
QY 1 MVIOGNSILLIIIFASSISTCRSNVIDDNLFKOVYDNIIEQEPFADFOAYLSYLSKNIE 60
DB 1 MVIOGNSILLIIIFASSISTCRSNVIDDNLFKOVYDNIIEQEPFADFOAYLSYLSKNIE 60
QY 61 SNNNIDKVDKNGIKIVINLSPGAKGDKGTYNIAFEQAMNEACSSRTPOVFPVKKNY 120
DB 61 SNNNIDKVDKNGIKIVINLSPGAKGDKGTYNIAFEQAMNEACSSRTPOVFPVKKNY 120
QY 121 LKQTFSPGCRSSISVKIFGSLSSKISDYKORRLWIAFDSYQNLVGGGGTINAGV 180
DB 121 LKQTFSPGCRSSISVKIFGSLSSKISDYKORRLWIAFDSYQNLVGGGGTINAGV 180
QY 181 WMPSSCKINKSLPRDAPALTFMCKNLKVNLLKSKNAQOIHIFESCTNVASNLMIN 240
DB 181 WMPSSCKINKSLPRDAPALTFMCKNLKVNLLKSKNAQOIHIFESCTNVASNLMIN 240
QY 241 ASAKSPMTDGVHVSNTQYIQISDTIIGTGDDCISIVSGSONVQATNITGPGHGISTGL 300
DB 241 ASAKSPMTDGVHVSNTQYIQISDTIIGTGDDCISIVSGSONVQATNITGPGHGISTGL 300
QY 301 GSGNSEAYVSNVTVAEAKIIGAENGRIKTWOGGSGOASNIFELNEMQDVYKPIIIDQN 360
DB 301 GSGNSEAYVSNVTVAEAKIIGAENGRIKTWOGGSGOASNIFELNEMQDVYKPIIIDQN 360
QY 361 YCDRVEPTCIQPSAVQVKNVYENIKGTSATKVAIKFDPSTNFPCEGIMININLVGSSG 420
DB 361 YCDRVEPTCIQPSAVQVKNVYENIKGTSATKVAIKFDPSTNFPCEGIMININLVGSSG 420
QY 421 KPESEATCKVHFNMAEHVPHCTSLFISDEALYNY 457
DB 421 KPESEATCKVHFNMAEHVPHCTSLFISDEALYNY 456
```

#### RESULT 12

US-09-791-537-49108  
Sequence 49108, Application US/09791537  
GENERAL INFORMATION:  
APPLICANT: Biomomix, Inc.  
APPLICANT: Debe, Derek  
TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS  
FILE REFERENCE: 261/210  
CURRENT FILING DATE: 2001-02-22  
CURRENT APPLICATION NUMBER: US/09/791.537  
NUMBER OF SEQ ID NOS: 153055  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 49108  
LENGTH: 467  
TYPE: PRT  
ORGANISM: Actinidia chinensis  
US-09-791-537-49108

Query Match 54.2%; Score 1295.5; DB 22; Length 467;  
Best Local Similarity 54.6%; Pred. No. 6.7e-120;  
Matches 257; Conservative 75; Mismatches 112; Indels 27; Gaps 7;

```
QY 1 MVIOGN--SILLIIIFASSISTCRSNVIDDNLFKOVYDNIIEQEPFADFOAYLSYLSKN 58
DB 1 MAQGRFPQFQVITTLTLPSTLIGTYSAVHD-----PPHDYHLE-FYGYDFKAYPSYITTI 55
QY 59 IESNNNIDKVDKNGI-----KVINLSPGAKGDKGTYNIAFEQAMNE 101
DB 59 GNDPFGSSKSHENGIFGKRVKVDYGDWRVLDASKYVNVDDFGAKGGRD-TRKAEKAKWA 114
QY 102 ACCSRTPGVFPVKKNYILKQITSGPCRSSISVYKIFGSLSSKISDY-KDRRLWIAF 160
DB 115 ACCSTSSAVTLVPR-KNYLVPRISFGPKCKGLTWQIYGTLEASDDSDYRKQDRHMLVF 173
```

```
QY 161 DSVQNLVYGGGGTINAGVWMPSSCKINKSLPRDAPALTFMCKNLKVNLLKSKNAQ 220
DB 174 DSVQNLVYGGGGTINAGVWMPSSCKINKSLPRDAPALTFMCKNLKVNLLKSKNAQ 223
QY 221 QIHIFESCTNVASNLMINASAKSPMTDGVHVSNTQYIQISDTIIGTGDDCISIVSGO 280
DB 234 QIHVSFDCVNVQASNLMTVAIPNSPMTDGIHTGTQNHISCVLTGTDDCISIVSGR 293
QY 294 KVRVNDITGPGHGISTGSGNSEAYVSNVTVAEAKIIGAENGRIKTWOGGSGOASN 340
DB 294 KVRVNDITGPGHGISTGSGNSEAYVSNVTVAEAKIIGAENGRIKTWOGGSGOASN 353
QY 341 IKFLNEMQDVYKPIIIDQNYCDRVEPTCIQPSAVQVKNVYENIKGTSATKVAIKFDCS 400
DB 354 IKFQNVEMHNVENPIIIDQNYCDRVEPTCIQPSAVQVKNVYENIKGTSATKVAIKFDCS 413
QY 401 TNPCEGIMININLVGSGKPESEATCKVHFNMAEHVPHCTSLFISDE 451
DB 414 KPEPCQGIYLEDVLEIEGGAATAKALCNVVELSEGTGVSPHCQEGEGEBE 464
```

#### RESULT 13

US-09-791-537-9525  
Sequence 9525, Application US/09791537  
GENERAL INFORMATION:  
APPLICANT: Biomomix, Inc.  
APPLICANT: Debe, Derek  
TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS  
FILE REFERENCE: 261/210  
CURRENT FILING DATE: 2001-02-22  
CURRENT APPLICATION NUMBER: US/09/791.537  
NUMBER OF SEQ ID NOS: 153055  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 9525  
LENGTH: 438  
TYPE: PRT  
ORGANISM: Arabidopsis thaliana  
US-09-791-537-9525

Query Match 49.6%; Score 1185; DB 22; Length 438;  
Best Local Similarity 51.9%; Pred. No. 7.2e-109;  
Matches 230; Conservative 76; Mismatches 121; Indels 16; Gaps 5;

```
QY 8 ILLIIIFASSISTCRSN---IDNLFQO---VYDNIIEQEPFADFOAYLSYLSKNIES 61
DB 5 ILLISVLLTFLPSCRSSYEPFNRDNLFWMSNVYETNRQHGHN-----TRNSHL 56
QY 62 NNNIDKVDKNGIKIVINLSPGAKGDKGTYNIAFEQAMNEACSSRTPOVFPVKKNY 121
DB 57 KNRHGVAPRRSPRSPFVNTFGAKAKAND-DSKAFKAMWAAACSTGIYIYAPKRRDYML 115
QY 122 KQTFSPGCRSSISVKIFGSLSSKISDYKORRLWIAFDSYQNLVGGGGTINAGV 181
DB 116 KAVTFSGPKCKSLIIFIRIEMENPSPDYKERRHMYIFEVANNLRVGGGRIDANGIIM 175
QY 182 WMPSSCKINKSLPRDAPALTFMCKNLKVNLLKSKNAQOIHIFESCTNVASNLMINA 241
DB 176 WPKSKINPOLCLAPLAVTVFECNNLRVSNIRLENAQOMHLTFQDCKNKVALMLWTS 235
QY 242 SAKSPMTDGVHVSNTQYIQISDTIIGTGDDCISIVSGSONVQATNITGPGHGISTGL 301
DB 236 PADSPMTDGIHVSQNLIDQSIVRTDDCISIVSGSENVRAICITGPGHGISTGL 295
QY 302 GSGNSEAYVSNVTVAEAKIIGAENGRIKTWOGGSGOASNIFELNEMQDVYKPIIIDQN 361
DB 296 EDNSAIVSNVTVAEAKIIGAENGRIKTWOGGSGOAKNIIFODIIMKNTNPIIIDQN 355
QY 362 CDRVEPTCIQPSAVQVKNVYENIKGTSATKVAIKFDPSTNFPCEGIMININLVGSSG 421
DB 356 CDRVEACPEQKSAVQVSNVLTAKNIQGTSSRPVIAVKFVCSNKIIPKGISQNVKLVDTQO 415
```

Qy 422 P-SEATCKNVHFNNAEHVTEPCT 443  
Db 416 DVSXKASCSNVKLDTRGVNSPLCT 438

RESULT 14  
US-09-791-537-83144  
Sequence 83144, Application US/09791537  
GENERAL INFORMATION:  
APPLICANT: Bionomix, Inc.  
APPLICANT: Debe, Derek  
APPLICANT: Danzer, Joseph  
TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS  
FILE REFERENCE: 261/210  
CURRENT APPLICATION NUMBER: US/09/791,537  
CURRENT FILING DATE: 2001-02-22  
NUMBER OF SEQ ID NOS: 153055  
SOFTWARE: Patentin version 3.0  
SEQ ID NO 83144  
LENGTH: 453  
TYPE: PRT  
ORGANISM: Persea americana  
US-09-791-537-83144

Query Match 48.2%; Score 1153; DB 22; Length 453;  
Best Local Similarity 55.6%; Pred. No. 1.2e-105;  
Matches 224; Conservative 66; Mismatches 104; Indels 8; Gaps 5;

Qy 28 DNLFFQVNDNLEQFAIDFQAYLSLKNIESNNNDKVDKNGIKVINVLSPGAKGDG 87  
Db 40 DQRAYPTFGPILD-EFSSIMGFEPILS--LERFNPVGPEPSPPTDISVDGFGAGDG 96  
Qy 88 KTYDNIAPFOANWACSSRTPVOFVVPKKNVYLKQITFSGPCRSSISVKIFGSLBASG 147  
Db 97 -TDDTAFAFEKAMKDACSSGSVL--IYPENKVIKQITFSGPCSSDLRVKIRGTIEASSD 153  
Qy 148 ISDY--KDRRLWIAPDSVONLVVGGGTTNGQVWPPSSCKINKSLPCRDAPFTALTFFN 205  
Db 154 QSDWVGHNKRKRWLEFEDISNLTLEGGTTNGGETWMDSSCKRKKSLPCASAPFTALTFRS 213  
Qy 206 CKNLKYNLKSXAQOIHKFESCTNVASNLIMNASAKSPNTDGVHVSNTQYIQISDTI 265  
Db 214 CKNLIVSDLSIKDSQKHLSPKCDVIANLWVTAPEHSPTDGIHITGTQRIHWNVS 273  
Qy 266 IGTGDDCISIVSGSQNVQATNITCGPHGISIGSLGSGNSEAYVNTVNEAKIIGAENG 325  
Db 274 IGTGDDCISIESGSKWVIATNITCGPHGISIGSLGDRNSEAHVSGVLVDGGLFPTTNG 333  
Qy 326 VRIKTVQSGSGQASNIKFLNVEMQVYPIIIDONYCDRVEPCIQFSAYQVKNVYENI 385  
Db 334 LRIKTVQSGSGSAKNIKFQNIIVHNVNTPIIIDQYICDSKDCPEQESAVKSNVAVYMI 393  
Qy 386 KGTSAATKVAIKFDCSTNFPCEGIIMENINLVESGKPESEATCKNV 430  
Db 394 RGTSAEVAIVKFDCKSSPCQGIILENINLVGNGKETTMSCSNI 438

RESULT 15  
US-09-791-537-75907  
Sequence 75907, Application US/09791537  
GENERAL INFORMATION:  
APPLICANT: Bionomix, Inc.  
APPLICANT: Debe, Derek  
APPLICANT: Danzer, Joseph  
TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS  
FILE REFERENCE: 261/210  
CURRENT APPLICATION NUMBER: US/09/791,537  
CURRENT FILING DATE: 2001-02-22  
NUMBER OF SEQ ID NOS: 153055  
SOFTWARE: Patentin version 3.0  
SEQ ID NO 75907

; LENGTH: 462  
; TYPE: PRT  
; ORGANISM: Persea americana  
US-09-791-537-75907

Query Match 47.8%; Score 1142; DB 22; Length 462;  
Best Local Similarity 55.3%; Pred. No. 1.6e-104;  
Matches 224; Conservative 67; Mismatches 106; Indels 8; Gaps 5;

Qy 28 DNLFFQVNDNLEQFAIDFQAYLSLKNIESNNNDKVDKNGIKVINVLSPGAKGDG 87  
Db 40 DQRAYPTFGPILD-EFSSIMGFEPILS--LERFNPVGPEPSPPTDISVDGFGAGDG 96  
Qy 88 KTYDNIAPFOANWACSSRTPVOFVVPKKNVYLKQITFSGPCRSSISVKIFGSLBASG 147  
Db 97 -TDDTAFAFEKAMKDACSSGSVL--IYPENKVIKQITFSGPCSSDLRVKIRGTIEASSD 153  
Qy 148 ISDY--KDRRLWIAPDSVONLVVGGGTTNGQVWPPSSCKINKSLPCRDAPFTALTFFN 205  
Db 154 QSDWVGHNKRKRWLEFEDISNLTLEGGTTNGGETWMDSSCKRKKSLPCASAPFTALTFRS 213  
Qy 206 CKNLKYNLKSXAQOIHKFESCTNVASNLIMNASAKSPNTDGVHVSNTQYIQISDTI 265  
Db 214 CKNLIVSDLSIKDSQKHLSPKCDVIANLWVTAPEHSPTDGIHITGTQRIHWNVS 273  
Qy 266 IGTGDDCISIVSGSQNVQATNITCGPHGISIGSLGSGNSEAYVNTVNEAKIIGAENG 325  
Db 274 IGTGDDCISIESGSKWVIATNITCGPHGISIGSLGDRNSEAHVSGVLVDGGLFPTTNG 333  
Qy 326 VRIKTVQSGSGQASNIKFLNVEMQVYPIIIDONYCDRVEPCIQFSAYQVKNVYENI 385  
Db 334 LRIKTVQSGSGSAKNIKFQNIIVHNVNTPIIIDQYICDSKDCPEQESAVKSNVAVYMI 393  
Qy 386 KGTSAATKVAIKFDCSTNFPCEGIIMENINLVESGKPESEATCKNV 430  
Db 394 RGTSAEVAIVKFDCKSSPCQGIIVGNINLVGNGKETTMSCSNI 438

Search completed: March 24, 2005, 07:15:02  
Job time: 487 secs

This Page Blank (uspio)

This Page Blank (uspio)

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Comugen Ltd.

OM protein - protein search, using sw model

Run on: March 24, 2005, 07:02:13 ; Search time 1256 Seconds  
(without alignments)  
120.472 Million cell updates/sec

Title: US-10-691-374-2  
Perfect score: 2390  
Sequence: 1 MVIGRNSILLIRIFASIS.....VTPHCTSLRISDEALVNY 457

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1407402 seqs, 331100923 residues

Total number of hits satisfying chosen parameters: 1407402

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications AA.\*  
1: /cgn2\_6/ptodata/2/pubppa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/2/pubppa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/2/pubppa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/2/pubppa/US07\_NEW\_PUB.pep.\*  
5: /cgn2\_6/ptodata/2/pubppa/PCTUS\_PUBCOMB.pep.\*  
6: /cgn2\_6/ptodata/2/pubppa/US10\_NEW\_PUB.pep.\*  
7: /cgn2\_6/ptodata/2/pubppa/US08\_PUBCOMB.pep.\*  
8: /cgn2\_6/ptodata/2/pubppa/US09\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/2/pubppa/US05\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/2/pubppa/US09C\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/2/pubppa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/2/pubppa/US09\_NEW\_PUB.pep.\*  
13: /cgn2\_6/ptodata/2/pubppa/US10\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/2/pubppa/US10\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/2/pubppa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/2/pubppa/US10D\_PUBCOMB.pep.\*  
17: /cgn2\_6/ptodata/2/pubppa/US11\_NEW\_PUB.pep.\*  
18: /cgn2\_6/ptodata/2/pubppa/US11\_NEW\_PUB.pep.\*  
19: /cgn2\_6/ptodata/2/pubppa/US60\_NEW\_PUB.pep.\*  
20: /cgn2\_6/ptodata/2/pubppa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description          |
|------------|-------|-------------|--------|-------|----------------------|
| 1          | 1113  | 46.6        | 433    | 13    | US-10-151-668-2      |
| 2          | 1103  | 46.2        | 433    | 16    | US-10-787-958-32     |
| 3          | 1011  | 42.3        | 393    | 16    | US-10-437-963-127090 |
| 4          | 1010  | 42.3        | 452    | 16    | US-10-437-963-127091 |
| 5          | 988.5 | 41.4        | 444    | 16    | US-10-437-963-131044 |
| 6          | 870   | 36.4        | 319    | 15    | US-10-425-114-69361  |
| 7          | 808   | 33.8        | 508    | 16    | US-10-437-963-154485 |
| 8          | 787.5 | 32.9        | 514    | 10    | US-09-847-208-69     |
| 9          | 785   | 32.8        | 534    | 16    | US-10-437-963-134257 |
| 10         | 739   | 30.9        | 485    | 16    | US-10-437-963-175508 |
| 11         | 734   | 30.7        | 485    | 16    | US-10-437-963-172689 |
| 12         | 728   | 30.5        | 443    | 15    | US-10-424-599-235443 |
| 13         | 728   | 30.5        | 456    | 15    | US-10-425-114-44707  |

|    |       |      |     |    |                      |                   |
|----|-------|------|-----|----|----------------------|-------------------|
| 14 | 724.5 | 30.3 | 398 | 15 | US-10-362-091-4      | Sequence 4, Appl1 |
| 15 | 723.5 | 30.3 | 490 | 16 | US-10-424-599-284649 | Sequence 284649,  |
| 16 | 701.5 | 29.4 | 402 | 15 | US-10-437-963-155040 | Sequence 155040,  |
| 17 | 700   | 29.3 | 573 | 15 | US-10-425-114-43413  | Sequence 43413, A |
| 18 | 699   | 29.2 | 496 | 15 | US-10-425-114-66151  | Sequence 66151, A |
| 19 | 682   | 28.5 | 571 | 16 | US-10-437-963-159100 | Sequence 159100,  |
| 20 | 681   | 28.5 | 737 | 16 | US-10-437-963-162350 | Sequence 162350,  |
| 21 | 678   | 28.4 | 457 | 16 | US-10-437-963-146166 | Sequence 146166,  |
| 22 | 666   | 27.9 | 393 | 16 | US-10-437-963-116114 | Sequence 116114,  |
| 23 | 664.5 | 27.8 | 415 | 15 | US-10-424-599-238512 | Sequence 238512,  |
| 24 | 660   | 27.6 | 422 | 15 | US-10-424-599-176320 | Sequence 176320,  |
| 25 | 653.5 | 27.3 | 385 | 15 | US-10-425-114-41984  | Sequence 41984, A |
| 26 | 653.5 | 27.3 | 411 | 15 | US-10-425-114-38472  | Sequence 38472, A |
| 27 | 653.5 | 27.3 | 415 | 15 | US-10-425-114-51531  | Sequence 51531, A |
| 28 | 649.5 | 27.2 | 449 | 16 | US-10-437-963-110174 | Sequence 110174,  |
| 29 | 643.5 | 26.9 | 602 | 16 | US-10-437-963-127089 | Sequence 127089,  |
| 30 | 642   | 26.9 | 320 | 15 | US-10-424-599-204806 | Sequence 204806,  |
| 31 | 641.5 | 26.8 | 544 | 16 | US-10-437-963-110166 | Sequence 110166,  |
| 32 | 635.5 | 26.6 | 539 | 15 | US-10-424-599-176466 | Sequence 176466,  |
| 33 | 632   | 26.4 | 386 | 15 | US-10-425-114-50981  | Sequence 50981, A |
| 34 | 630   | 26.4 | 410 | 15 | US-10-425-114-38475  | Sequence 38475, A |
| 35 | 630   | 26.4 | 414 | 15 | US-10-425-114-62958  | Sequence 62958, A |
| 36 | 630   | 26.4 | 414 | 15 | US-10-425-114-64531  | Sequence 64531, A |
| 37 | 627   | 26.2 | 407 | 15 | US-10-425-114-65565  | Sequence 65565, A |
| 38 | 627   | 26.2 | 410 | 15 | US-10-425-114-46486  | Sequence 46486, A |
| 39 | 627   | 26.2 | 412 | 15 | US-10-425-114-46801  | Sequence 46801, A |
| 40 | 627   | 26.2 | 413 | 15 | US-10-425-114-52982  | Sequence 52982, A |
| 41 | 627   | 26.2 | 413 | 15 | US-10-425-114-56708  | Sequence 56708, A |
| 42 | 627   | 26.2 | 415 | 15 | US-10-425-114-46800  | Sequence 46800, A |
| 43 | 627   | 26.2 | 415 | 15 | US-10-425-114-54673  | Sequence 54673, A |
| 44 | 627   | 26.2 | 415 | 15 | US-10-425-114-64510  | Sequence 64510, A |
| 45 | 627   | 26.2 | 416 | 15 | US-10-425-114-65564  | Sequence 65564, A |

## ALIGNMENTS

RESULT 1  
US-10-151-668-2  
Sequence 2, Application US/10151668  
Publication No. US20020184660A1  
GENERAL INFORMATION:  
APPLICANT: ULVSKOV, Peter  
APPLICANT: CHILD, Robin  
APPLICANT: VAN ONCKELIN, Henri  
APPLICANT: PRINSEN, Eib  
APPLICANT: BORKHARDT, Bernard  
APPLICANT: SANDER, Lillil  
APPLICANT: PETERSEN, Morten  
APPLICANT: BUNDEARD, Poulsen, Gert  
TITLE OF INVENTION: Seed Shattering  
FILE REFERENCE: 2121-0138P  
CURRENT APPLICATION NUMBER: US/10/151,668  
CURRENT FILING DATE: 2002-05-21  
PRIOR APPLICATION NUMBER: US/09/051,239  
PRIOR FILING DATE: 1998-09-28  
PRIOR APPLICATION NUMBER: PCT/EP96/04313  
PRIOR FILING DATE: 1996-10-04  
PRIOR APPLICATION NUMBER: EP 95 402241.4  
PRIOR FILING DATE: 1995-10-06  
PRIOR APPLICATION NUMBER: EP 95 203328.0  
PRIOR FILING DATE: 1995-12-08  
NUMBER OF SEQ ID NOS: 14  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 2  
LENGTH: 433  
TYPE: PRT  
ORGANISM: Brassica napus  
FEATURE:  
OTHER INFORMATION: Strain cv. Topaz.  
US-10-151-668-2

```

Query Match      46.6%; Score 1113; DB 13; Length 433;
Best Local Similarity 49.9%; Pred. No. 5.6e-91;
Matches 223; Conservative 69; Mismatches 121; Indels 34; Gaps 8;

QY 7 SILLIIIFASSISTCRSNVIDNLFKQVYDNLIEQFAHDFQAVYSTKSNINND 66
DB 9 AVELCVLMLACCOALLSNV-DDG-----YGHEDG---SFESDLIKLAND 51
QY 67 KV-----DKNGIKYNLISFGAKGDKTYNTNIAFEQANMEACSSRTPVQFVVPKXN 118
DB 52 DVLTKSDDPTRETSSTVSISNFGAKGDKTDPTQAFKAMKKAQSTNGVTTFLLPKGKT 111
QY 119 YLLKQTFSGPCRSSISVVKIFGSLKSIQSDY-KDRRLMIAFDSYQNLVGGG--GTIN 175
DB 112 YLLKSIREFRGPCKSLRSFQILGTLASSTRSDYSNDKNHMLLEDVNNLSIGSGAGLYD 171
QY 176 GNGQVWPPSSCKINSLPCRDAPALTTFMCKKLYNNLSKXNAQOIHKFESCTNVAS 235
DB 172 GNGNIMWQNSCKIDSKPCTKAPTALTYLNKLNKLNKLRVNAQOIQISIEKCNVGVK 231
QY 236 NLMINASAKSPNTDGVHVSNTQYIQTSDTIIGTGDCTISVSGSONVQATNITCGPHGI 295
DB 232 NYKTFAPGDSPTDGHIVATKNIRISNDISTGDDCTISIEGSONVQINDLTCGPHGI 291
QY 296 STIGSLGSGSEAYSVNTVNEAKIIGAENGVRIKTWQSGGSAQSNIKFLANVEMOVKYP 355
DB 292 STIGSLGDDSKAYVSGIDVDTALSETDNGVRIKTYQSGGTAKNIKFQNRIMNVKNPI 351
QY 356 IIDQNYCDRVPCIIQOFSAVOYQNVYENIKGTSTATKYAIKFDSTNPPCGIIMENINL 415
DB 352 IIDQNYCDK-DKCEQOESAVOVNNAVYQNIQTSATDVAIMFNCSSVKYPCQGIYLENVNI 410
QY 416 VQESKPSSEATCKNVHFNNAEHVTPHC 442
DB 411 KG-----GKASCENVNVKDKGTVSPKC 432

RESULT 2
US-10-787-958-32
; Sequence 32, Application US/10787958
; Publication NO. US20040154053A1
; GENERAL INFORMATION:
; APPLICANT: Wyatt, Paul
; APPLICANT: Roberts, Jeremy A.
; APPLICANT: Whitefaw, Catherine
; TITLE OF INVENTION: Signal Transduction Protein Involved in Plant Dehiscence
; FILE REFERENCE: 0623, 089000
; CURRENT APPLICATION NUMBER: US/10/787,958
; PRIOR FILING DATE: 2004-02-27
; PRIOR APPLICATION NUMBER: GB9806113.8
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 32
; LENGTH: 433
; TYPE: PRT
; ORGANISM: Brassica napus
US-10-787-958-32

Query Match      46.2%; Score 1103; DB 16; Length 433;
Best Local Similarity 49.8%; Pred. No. 4.4e-90;
Matches 221; Conservative 70; Mismatches 125; Indels 28; Gaps 7;

QY 7 SILLIIIFASSISTCRSNVIDNLFKQVYDNLIEQFAHDFQAVYSTKSNIES 61
DB 9 AIFLCVILMLACCOALLSNVDDYGHEDGSFET--DSLILKLANDDVLTLKSSDRPTTES 66
QY 62 NNNIDVNDKGIKYNLISFGAKGDKTYNTNIAFEQANMEACSSRTPVQFVVPKXNYL 121
DB 67 S-----TVSVSNFGAKGDKTDPTQAFKAMKKAQSTNGVTTFLLPKGKTYL 114
QY 122 KQITSGPCRSSISVVKIFGSLKSIQSDY-KDRRLMIAFDSYQNLVGGG--GTINGG 178
DB 122 KQITSGPCRSSISVVKIFGSLKSIQSDY-KDRRLMIAFDSYQNLVGGG--GTINGG 178

```

```

DB 115 KSIRFRGPCKSLRSFQILGTLASSTRSDYSNDKNHMLLEDVNNLSIDGSGAGLYDNG 174
QY 179 QVWPPSSCKINKSLPCRDAPALTTFMCKKLYNNLSKXNAQOIHKFESCTNVASNL 238
DB 175 KIMWQNSCKIDSKPCTKAPTALTYLNKLNKLNKLRVNAQOIQISIEKCNVGVK 234
QY 239 INSAKSPNTDGVHVSNTQYIQTSDTIIGTGDCTISVSGSONVQATNITCGPHGISIG 298
DB 235 ITAFGDSPTDGHIVATKNIRISNDISTGDDCTISIEGSONVQINDLTCGPHGISIG 294
QY 299 STIGSGSEAYSVNTVNEAKIIGAENGVRIKTWQSGGSAQSNIKFLANVEMOVKYP 358
DB 295 SLGDNSKAYVSGINDVDTALSETDNGVRIKTYQSGGTAKNIKFQNIEMDNVKNPIIID 354
QY 359 QNYCDRVPCIIQOFSAVOYQNVYENIKGTSTATKYAIKFDCTNPPCGIIMENINLGE 418
DB 355 QNYCDK-DKCEQOESAVOVNNAVYQNIQTSATDVAIMFNCSSVKYPCQGIYLENVNIG- 412
QY 419 SGKPSSEATCKNVHFNNAEHVTPHC 442
DB 413 ---GKASCENVNVKDKGTVSPKC 432

RESULT 3
US-10-437-963-127090
; Sequence 127090, Application US/10437963
; Publication NO. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; PRIOR FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 127090
; LENGTH: 393
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_29575C.1.pep
US-10-437-963-127090

Query Match      42.3%; Score 1011; DB 16; Length 393;
Best Local Similarity 53.3%; Pred. No. 6.9e-82;
Matches 187; Conservative 57; Mismatches 105; Indels 2; Gaps 1;

QY 72 GIKYNLISFGAKGDKTYNTNIAFEQANMEACSSRTPVQFVVPKXNYLLKQITSGPCR 131
DB 27 GSNVFSIQSYGAHGDHDPDKALGDTWAAACSSAKPAVLLPKGKXLYIKTTISGPK 86
QY 132 SSIIVKIFGSLKSIQSDYKDR--RLMIAFDSYQNLVGGGCTINGGNGQVWPPSSCKIN 189
DB 87 SSIILWQNSLVASPRSDMSKETIRHMLISGVTLVYGGGTLIDGNGKIWMQNSCKTN 146
QY 190 KSLPCRDAPALTTFMCKKLYNNLSKXNAQOIHKFESCTNVASNLMINASAKSPNT 249
DB 147 SKLPCTEAPALTTFMCKKLYNNLSKXNAQOIQISVEDCTDVMVSRSLITAPETAPNTD 206
QY 250 GNVHVSNTQYIQTSDTIIGTGDCTISVSGSONVQATNITCGPHGISIGSLGSGSEAYV 309
DB 207 GHITRSRVEVTDCKIKGDDCMGSIEDGTENLHVKNVCGPHGISIGSLGDNSEAHV 266
QY 310 SNVTVEAKIIGAENGVRIKTWQSGGSAQSNIKFLANVEMOVKYP 369
DB 267 NNVTVDNVRILYGTANGAIKTWQSGGSAQSNIVFQNMVNDVNMPIIIDQNYCDSSTPCK 326

```



APPLICANT: Barbazuk, Brad  
TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With  
FILE REFERENCE: 38-21(53221)B  
CURRENT APPLICATION NUMBER: US/10/437,963  
NUMBER OF SEQ ID NOS: 204966  
SEQ ID NO 131044  
LENGTH: 444  
TYPE: PRT  
ORGANISM: Oryza sativa  
FEATURE:  
OTHER INFORMATION: Clone ID: PAT\_MRT4530\_33148C.1.pep  
US-10-437-963-131044

Query Match 41.4%; Score 988.5; DB 16; Length 444;  
Best Local Similarity 49.5%; Pred. No. 8,6e-80;  
Matches 186; Conservative 73; Mismatches 110; Indels 7; Gaps 5;

Oy 75 VINVLSFGAKGDKTYDNIAPQAMNEACSSRTPVGFVPKKNLYLKQITFSGPCRSSTI 134  
:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|  
Db VFSLDSYGAHGGERDDTALARAWSAACASAAPAVLVBESRSYLRLQVTLSGPCESTI 127  
68 SVKIFGSLEASKSIDKYD -RRLWIAFPSVQLVYGGGTINGNGVWWPPSSCKINKSL 192  
Oy 135 SVKIFGSLEASKSIDKYD -RRLWIAFPSVQLVYGGGTINGNGVWWPPSSCKINKSL 192  
128 KLMVGTLVAASPDMNMSNRNYLVIRGVGLAVAGGGDTIDNGEGMENSCKINRAL 187  
Oy 193 PCRDAPTALTFFNCCKNLKYNLNKSNAQQIHIFECSCTNVVASNLMAASAKSPNTDGVH 252  
188 PCKGAPTALSFITCDNLNSGLKMNVSQDIHMSVECTGYELHLISASACTSNTTGIGH 247  
Oy 253 VSNTOYIOISDTTIISTGDCCISIVSGSQNVQATNITCGPHGISIGSLGSGNSEAYYSNV 312  
248 ITHSKNVQVSDCTIKTGDDCVSIEDGTGHLHTVRLVCGRPHGISIGSLGDNSPAAYSVDI 307  
Oy 313 TVNEAKTIGAENGVARIKTVQGSGGQASNKFLENVEDVRYPIIIDONYCDRAEPC-IQQ 371  
Db 308 FIDTVALYLTNGAARIKTWQSGGYAKDIVFOVMWNVSVPNPIIIDONYCDSAKKETQE 367  
Oy 372 FSADVOKNVVENIGTSATKVAIKPDCSTNPFCEGIEMENINL--VGEAGKPSEATCKN 429  
Db 368 GSAVEISNVVFNLGITTSKSAITLNCNKVPCYDISLDINDENVDDNG-KTGSTCN 426  
Oy 430 VFNNAEHVTPH-CTS 444  
Db 427 AKMRKSGTVVPQPCTS 442

RESULT 6  
US-10-425-114-69361  
Sequence 69361, Application US/10425114  
Publication No. US20040034888A1  
GENERAL INFORMATION:  
APPLICANT: Liu, Jinding  
APPLICANT: Zhou, Yihua  
APPLICANT: Kovacic, David K.  
APPLICANT: Screen, Steven E  
APPLICANT: Tabaska, Jack B  
APPLICANT: Cao, Yongwei  
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
FILE REFERENCE: 38-21(53313)B  
CURRENT APPLICATION NUMBER: US/10/425,114  
CURRENT FILING DATE: 2003-04-28  
NUMBER OF SEQ ID NOS: 73128  
SEQ ID NO 69361  
LENGTH: 319  
TYPE: PRT  
ORGANISM: Zea mays  
FEATURE:  
OTHER INFORMATION: Clone ID: UC-ZMLJB73002C03\_FLI.pep  
US-10-425-114-69361

```

Query Match      36.4%; Score 870; DB 15; Length 319;
Best Local Similarity 50.8%; Pred. No. 2.3e-63;
Matches 162; Conservative 64; Mismatches 89; Indels 4; Gaps 3;

QY 132 SSISVKIFGSLSEASSKISDY--KDRRLMIAFDSVQNLVVGSGGTINGNGQVWMPSSCKIN 189
DB 1 SSVTLVYKGLTVASPNRADWSNDNRHMIYFRSIDLKTVNGGAILONGNEKMWPHSCKIN 60
QY 190 KSLPCRDAPFALTATFWNCKNLKVNVLKSKNAOQHIKFESECTNVVNASLMINASAKSPND 249
DB 61 KALPCKEAPFATLSHYCVDLKVEBKLKVSQQLHMSVESANLRLSTIATGTSPTND 120
QY 250 GVHNSNTQYIISDTIIGTDDCISIVSGSONVQATNITCGPHGISIGISGSEAV 309
DB 121 GIHTSKDVRVTDCKITGDDCMSIENGTHNLHVSKNVCGPHGISISLSDDDNSRAEV 180
QY 310 SNTVNEAKIIGENGRIKTMOGSGGQSNIKFELAVEMQDYKYPITIDONYCDREVERCI 369
DB 181 SGITISVOLHGTTHNGARIKTKYQSGGYAKDITFQNMVWYDVKNPIIIDONYCDKXPCG 240
QY 370 QQPSAVQVKNVYENIKGTSATKVAIKFDCSTNFPCEGIMENINLVGESGK-PSEATCK 428
DB 241 EOSEAVQVSGVVFKNIRGTTSTKDAIKMNCSENVFCQGITLQNLKMQDGKNTRESTCQ 300
QY 429 NVHFNAEHVTPH-CPSLE 446
DB 301 NAKWTEFGTVRPOPCIAIK 319

RESULT 7
US-10-437-963-154485
; Sequence 154485, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 154485
; LENGTH: 508
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_54340C.1.pep
US-10-437-963-154485

Query Match      33.8%; Score 808; DB 16; Length 508;
Best Local Similarity 36.4%; Pred. No. 1.6e-63;
Matches 173; Conservative 91; Mismatches 167; Indels 44; Gaps 7;

QY 8 ILLIITPASTISTCSNVIDNLFKQVYDNLLEQFAHDFQAVY-----SY 54
DB 18 VLVALMIVMAAKTASGGGDDGAGGDH-----HDHQQFKLMNGRGADAKEDY 69
QY 55 LSKNIESNNIDKVDKN-----GIKVINLISGAKGDDGTYNINAFQ 97
DB 70 LMMDDDDDDDEDEDEEBAEQVMAAKCRPPRAGRNVVNDVSGAAGDGSDDTEAFIN 129
QY 98 ANNEACSSRTPVQVVPKKNYILKQITFSGPCSSISVVKIFGSLSEASSKISDY--KDRR 155
DB 130 AKKACSLNNAV-FLVPGRRYKVGAAFIQPCRMKMIQIGITIVAPDEPSEMDPASPR 188

```

```

QY 156 LMIAFDSVQNLVVGSGGTINGNGQVWMPSSCKINKSLPCRDAPFALTATFWNCKNLKVNVLK 215
DB 189 LMLFSLGIAKARIQGGGLDSSGSKWMANCKIDRSPPCKGAPFALTIDS CRGVSNRLR 248
QY 216 SKNAOQHIKFESECTNVVNASLMINASAKSPNDGVHNSNTQYIISDTIIGTDDCISI 275
DB 249 LQNAQOMHLTVRSRSDVRLASVRVDSPEPNDGIHVADSTAVTIOGCRINATGDDCISI 308
QY 276 VSGSONVQATNITCGPHGISISLSSGSEAVSVTVNEAKIIGENGRIKTMOGGS 335
DB 309 SNGSFVVRMDIDCGPEHGISTISLGQGAFAVVDGVSIDGARVARAQGVRIKTMOGGA 368
QY 336 GQASNIKFLVEMQDYKYPITIDONYCDREVERCIOQFSAVQVKNVYENIKGTSATKVAI 395
DB 369 GYVRNVFAVVRVVDGVDHPIVIDQFYCDATRCRKRNTSVKVSQVVFNTGTARARAI 428
QY 396 KFCSTNFPCEGIMENINLVGESGKPSAECKN--VHFNAEHVTPHCTSLEIS 448
DB 429 KLASDAVPCVGIASDIDLREDDGGVEQVTCNCGMFGDDG-RVSPAADCLRTS 482

RESULT 8
US-09-847-208-69
; Sequence 69, Application US/09847208
; Publication No. US20030082190A1
; GENERAL INFORMATION:
; APPLICANT: Saxon, Andrew
; APPLICANT: Zhang, Ke
; APPLICANT: Zhu, Daocheng
; TITLE OF INVENTION: FUSION MOLECULES AND TREATMENT OF
; FILE REFERENCE: UCE67.002A
; CURRENT APPLICATION NUMBER: US/09/847,208
; NUMBER OF SEQ ID NOS: 177
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 69
; LENGTH: 514
; TYPE: PRT
; ORGANISM: Cryptomeria japonica (Japanese cedar)
US-09-847-208-69

Query Match      32.9%; Score 787.5; DB 10; Length 514;
Best Local Similarity 38.5%; Pred. No. 1.1e-61;
Matches 173; Conservative 78; Mismatches 163; Indels 35; Gaps 9;

QY 8 ILLIITPASTISTCSNVIDNLFKQVYDNLLEQFAHDFQAVYLSKNIENNNIDK 67
DB 12 VAMQLIIMAAA-----EDQSAQIMLDSIDQ-----YLR-----RSNRSILRK 47
QY 68 VD--KNGIKVINLISFGAKGDKTYDNIAPFOANNEACSSRTPVQVVPKKNYILKQI 124
DB 48 VHSRHDALNIFVEYKYGAVGDKHDCTEAFSTAWQAC-KKPSAMLLVPNGKKEVVNNL 106
QY 125 TEGSPCRSSISVYKIFGSLSEASSKISDYKDRRLMIAFDSVQNLVVGSGGTINGNGQVWMP 184
DB 107 FFGPQOPHFTFKVGIILAAQONPASMKNRNIWLFALTLGTTLMGKIVIGQGGKQMMAG 166
QY 185 SCK-INKSLPC--RDAPFALTATFWNCKNLKVNVLKSKNAOQHIKFESECTNVVNASLMINA 241
DB 167 QCKWNGRIGICNDRRPFAIKFDFSTGLIQLKLMNSPEHVLVFGNCEGVKIIGISITA 226
QY 242 SAKSPNDGVHNSNTQYIISDTIIGTDDCISIVSGSONVQATNITCGPHGISISIG 301
DB 227 PRDSPNTDGIIDIFAKNPHLQKNITIGTDDCVALTGSSSNVIEDLLCGPHGISISIG 286
QY 302 SGNSEAVSVTVNEAKIIGENGRIKTMOGSGGQSNIKFELAVEMQDYKYPITIDONY 361
DB 287 RENSRAEVSIVHVGAKFIDQNGLRIRIKTMOGSGMASHIIEYVENIMENSEPILINQFY 346
QY 362 CDRVERPCIQPSAVQVKNVYENIKGTSATKVAIKPDCSTNFPCEGIMENINLVGESGK 421
DB 347 CTSASACQNGRSANVQIDQVYIKNIRGTSATRAAIQLKCSDSMPCKODIKLSDISLKULTSGK 406

```



OTHER INFORMATION: Clone ID: PAT\_MRT4530\_70800C.1.pep  
US-10-437-963-172689

Query Match  
Best Local Similarity 30.7%; Score 734; DB 16; Length 503;  
Matches 153; Conservative 66; Mismatches 147; Indels 10; Gaps 6;

QY 78 VLSFGAKGDKTYDNIATFEQAMNEACSSRTPVQFVVPKKNKYLKQITFSGPCRSISIVK 137  
DB 114 VXDFFGAAGDVDDTDALKTAMDTCADDGAGVULAAAGRSFLHTTTFGTGCGSVTLQ 173  
QY 138 IFSGLSEASKISDY--KDRRLMIAFDSVQNLVVGSGGTINGQVWVWSSCKINKS---- 191  
DB 174 VDGTVAPSEPAITPANNKRWMLVFRADGVSLVAGAGIDGKGKMDLCPKHKKGNTN 233  
QY 197 LPCRAPALTFWNCNKLKVNLSKNAQOIHIFESCTNVVASNLMINASASPNTDGV 251  
DB 234 GPC-DSPVAMRFALSNNTVTRGLKQNSPEHFHFDGNGVRVDDGLSISPLSPMTDGI 292  
QY 252 HVSNTQYIQISDTIIGGDDCISIVSGSNVQATNITTCGPHGISISIGLSGNSSEAYVEN 311  
DB 293 HVENISDVLINTVVSNGDCVSIAGCTLNVHLENTVCGPHGISISIGLCKAGTACVAN 352  
QY 312 VTVNEAKIENGVRIRKTVQSGSGQASNIKFLNVEODVKYPIIIDQNYCDRVEPTQQ 371  
DB 353 VTVNNAVIRHSDNGVRIRKTVQSGSGSAVAFENVMADAVRNPFIIDQYCC-LSKSCENE 411  
QY 372 FSAVOVKVYVENIKGTSATK-VAIKFDCSTNPFCEGIMENINLVGSGKP-SEATCKN 429  
DB 412 TTVAVNGVSYSGIRGTIVRGPPIHFQGSDAVPCTNITLSDVELLPASGDTVEDEPCMN 471  
QY 430 VHRNNAEHVTPHCTSL 445  
DB 472 VYGNMAATPTVPVSVCL 487

## RESULT 12

US-10-424-599-235443  
Sequence 235443, Application US/10424599  
Publication No. US20040031072A1  
GENERAL INFORMATION:  
APPLICANT: La Rosa Thomas J  
APPLICANT: Kovalic David K  
APPLICANT: Zhou Yihua  
APPLICANT: Cao Yongwei  
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
FILE REFERENCE: 38-21(53223)B  
CURRENT APPLICATION NUMBER: US/10/424,599  
CURRENT FILING DATE: 2003-04-28  
NUMBER OF SEQ ID NOS: 285684  
SEQ ID NO 235443  
LENGTH: 443  
TYPE: PRT  
ORGANISM: Glycine max  
FEATURE:  
OTHER INFORMATION: Clone ID: PAT\_MRT3847\_54623C.1.pep  
US-10-424-599-235443

Query Match  
Best Local Similarity 30.5%; Score 728; DB 15; Length 443;  
Matches 158; Conservative 51; Mismatches 139; Indels 20; Gaps 10;

QY 77 NVLSFAKDDGKTYDNIATFEQAMNEACSSRTPVQFVVPKKNKYLKQITFSGP-CRSIS 135  
DB 47 NVLDYGAKDGHADDTKAFEDMAAACVGESEST-MVVPSSGVFLVKFISFGPCEPNIV 105  
QY 136 VIFGSLSEASKISDYKDRRL-WIAFDSVQNLVVGSGGTINGQVWV-----PSSC 186  
DB 106 FOLDGKIAPTSSEMGSGTLOMLEFSKLTNTITRGKGVIDGQSVWMNNDSPTYNPTVE 165  
QY 167 KI--NKSIPCRDAPALTFFWNCNKLKVNLSKNAQOIHIFESCTNVVASNLMINASAK 244

DB

166 MLESNGRLP-STKPTALRFYSGDGVTVGTITQNSQOQTHLKPFSCNTNVQVSGISVSSPED 224

QY

245 SPNTDGVHVSNTQYIQISDTIIGGDDCISIVSGSNVQATNITTCGPHGISISIGLSGN 304

DB

225 SPNTDGIHLQNSQNVVYSSITLACGDDCVSIQGCSDIYHNVNCCPGHGISISIGLSREN 284

QY

305 SEAYSNVTVNEAKIENGVRIRKTVQSGSGQASNIKFLNVEODVKYPIIIDQNYCDR 364

DB

295 TKACVRNVTVRVDVTIQTTLTGVRIRKTVQSGSGSVQINFSNVQVSGVQTPISIDQYCDG 344

QY

365 VEPICQFSAVOVKVYVENIKGTSATKVAIKFDCSTNPFCEGIMENINLVGSGKPS 424

DB

345 GR-CRNSSSAVAVSGIHVYVVKGT-YTKPITFACSDNLPSCGTLTDTIOL--ESAQ--E 398

QY

425 ATCKNVHF 432

DB

399 TKNSNVPE 406

## RESULT 13

US-10-425-114-44707  
Sequence 44707, Application US/10425114  
Publication No. US20040034888A1  
GENERAL INFORMATION:  
APPLICANT: Liu, Jingdong  
APPLICANT: Zhou, Yihua  
APPLICANT: Kovalic, David K.  
APPLICANT: Screen, Steven E  
APPLICANT: Tabaska, Jack E  
APPLICANT: Cao, Yongwei  
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
FILE REFERENCE: 38-21(5313)B  
CURRENT APPLICATION NUMBER: US/10/425,114  
CURRENT FILING DATE: 2003-04-28  
NUMBER OF SEQ ID NOS: 73128  
SEQ ID NO 44707  
LENGTH: 456  
TYPE: PRT  
ORGANISM: Glycine max  
FEATURE:  
OTHER INFORMATION: Clone ID: 700847226\_FLI.pep  
US-10-425-114-44707

Query Match  
Best Local Similarity 30.5%; Score 728; DB 15; Length 456;  
Matches 158; Conservative 51; Mismatches 139; Indels 20; Gaps 10;

QY 77 NVLSFAKDDGKTYDNIATFEQAMNEACSSRTPVQFVVPKKNKYLKQITFSGP-CRSIS 135  
DB 60 NVLDYGAKDGHADDTKAFEDMAAACVGESEST-MVVPSSGVFLVKFISFGPCEPNIV 118  
QY 136 VIFGSLSEASKISDYKDRRL-WIAFDSVQNLVVGSGGTINGQVWV-----PSSC 186  
DB 119 FOLDGKIAPTSSEMGSGTLOMLEFSKLTNTITRGKGVIDGQSVWMNNDSPTYNPTVE 178  
QY 187 KI--NKSIPCRDAPALTFFWNCNKLKVNLSKNAQOIHIFESCTNVVASNLMINASAK 244  
DB 179 MLESNGRLP-STKPTALRFYSGDGVTVGTITQNSQOQTHLKPFSCNTNVQVSGISVSSPED 237  
QY 245 SPNTDGVHVSNTQYIQISDTIIGGDDCISIVSGSNVQATNITTCGPHGISISIGLSGN 304  
DB 238 SPNTDGIHLQNSQNVVYSSITLACGDDCVSIQGCSDIYHNVNCCPGHGISISIGLSREN 297  
QY 305 SEAYSNVTVNEAKIENGVRIRKTVQSGSGQASNIKFLNVEODVKYPIIIDQNYCDR 364  
DB 298 TKACVRNVTVRVDVTIQTTLTGVRIRKTVQSGSGSVQINFSNVQVSGVQTPISIDQYCDG 357  
QY 365 VEPICQFSAVOVKVYVENIKGTSATKVAIKFDCSTNPFCEGIMENINLVGSGKPS 424  
DB 358 GR-CRNSSSAVAVSGIHVYVVKGT-YTKPITFACSDNLPSCGTLTDTIOL--ESAQ--E 411  
QY 425 ATCKNVHF 432

Db 412 TKNSVVF 419

RESULT 14

US-10-362-091-4

Sequence 4, Application US/10362091

Publication No. US20040049809A1

GENERAL INFORMATION:

APPLICANT: Instituto de Ciencia Aplicada e Tecnologia (ICAT)

TITLE OF INVENTION: Pear genes coding for b-Galactosidase, Pectin Methylsterase,

TITLE OF INVENTION: Polylacturonase, Expansin and their use.

FILE REFERENCE: none

CURRENT APPLICATION NUMBER: US/10/362,091

PRIOR FILING DATE: 2003-02-12

PRIOR APPLICATION NUMBER: PT 102511 C

NUMBER OF SEQ ID NOS: 31

SOFTWARE: Patentin version 3.1

SEQ ID NO 4

LENGTH: 398

TYPE: PRT

ORGANISM: Pyrus communis

US-10-362-091-4

Query Match 30.3%; Score 724.5; DB 15; Length 398;

Best Local Similarity 41.4%; Pred. No. 3.5e-56;

Matches 158; Conservative 64; Mismatches 145; Indels 15; Gaps 10;

64 NIDKVDKNGIKVNLVSPGAKGDKTYDNIAPEQANNEACSSRTPVQFVVPKKNVLLKQ 123

23 SINVDAAV-TFSVSLGAKADGSTDTKAFLSAMNACASVNPAYIVPAGR-FLGN 80

124 ITFSGRC-SSISVKIFGSLASAKISDYK---DRRLMIFDSVQNLVVGCGGTINGNQ 179

81 AVFSGCKNNAIFRIAGTLVAP---SDYVIGNAGWMLFQHVNGVTI-SGCVLDGQGT 136

180 VVWPSGCKINKSLPCRDAPALTTFWNCNKLKNNLKSNAQOIHIFESCTNVVANSIMLMI 239

137 GLM-DCK-SGSGSCSGAGATTLTFSNNVNVVSGLLSINQMPHIYVNGCONKMGCVK 193

240 NAAKSPNTDGVVNTQYIOISDTIIGTDDCISIVSGSONVQATNITGPGHGISIGS 299

194 NAGNBPNTGIVHVGSSGVTILDSKISTGDDCVSGPPTNMLEVAVGPGHGISIGS 253

300 LGGSGEAVYVNTVEAKIIGAENGVRITKW-QGSGQASNIKFLNVMDQVKYPIIID 358

254 LGHQOEAGVQNTVKTFTGTENGVRIKSGRPSTGTFARSILFQHIWMTNVQNPIVID 313

359 QNVCDVEPCIQPSAVQVKNVYENKTSATKVAIKPDCSTNPCEGIIIMENINLVE 418

314 QNTCPNDKCGPGQASGVKSDVTYQDIGHTSATBEVAVKFDSSMYPCNJRLODVKLT-Y 372

419 SGKPSKATCKNVHFNNAHYTP 440

373 NQAASASCIHAGTTAGTVQ 394

RESULT 15

US-10-424-599-284649

Sequence 284649, Application US/10424599

Publication No. US20040031072A1

GENERAL INFORMATION:

APPLICANT: La Rosa Thomas J

APPLICANT: Kovalic David K

APPLICANT: Zhou Yihua

APPLICANT: Cao Yongwei

TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With

TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement

FILE REFERENCE: 38-21(35223)B

CURRENT APPLICATION NUMBER: US/10/424,599

CURRENT FILING DATE: 2003-04-28

NUMBER OF SEQ ID NOS: 285684

SEQ ID NO 284649

LENGTH: 492

TYPE: PRT

ORGANISM: Glycine max

FEATURE:

OTHER INFORMATION: Clone ID: PAT\_MRT3847\_99063C.1.pep

US-10-424-599-284649

Query Match 30.3%; Score 723.5; DB 15; Length 492;

Best Local Similarity 39.6%; Pred. No. 5.9e-56;

Matches 155; Conservative 64; Mismatches 157; Indels 15; Gaps 7;

75 VINLSPGAKGDKTYDNIAPEQANNEACSSRTPVQFVVPKKNVLLKQITFSGPCRSST 134

96 VFDVRSFGAVGDCADDTAFRAAWKAAVDSGI-VLAPENYSFKITSTIFSGPCPKPL 154

135 SVKIFGSLASAKISDYK---RRLLMIFDSVQNLVVGCGGTINGNQVWVWPSCKINK- 130

155 VFQVDTGLMAPDGPNSWPEADSENQWLVFYRLDQMTLNGTITEGNGDKWMDLPCKPHRG 214

191 -----SLPCRDAPALTTFWNCNKLKNNLKSNAQOIHIFESCTNVVANSIMINASAK 244

215 PNGKTLSPGK-SPAMIRPFMSNKLKVKIKIYNSFQPHMIFNGCGVLIDKLSISSPKL 273

245 SPNTDGVVNTQYIOISDTIIGTDDCISIVSGSONVQATNITGPGHGISIGSLGSGN 304

274 SPNTDGIHENSXYVGIYNSMISNGDDCISIGPSSNVDAJLTGPGSHGISIGSLGVHN 333

305 SEAYSVNTVNEAKIIGAENGVRITKWQSGGQASNIKFLNVMDQVKYPIIIDQNYCDR 364

334 SQACVSNLTVRDSIIRSDNGIRIKTWQGGMGSSVSLRFENIOMENVGNCIIIDQYYCLS 393

365 VEPICQPSAVQVKNVYENKTSATKVAIKPDCSTNPCEGIIIMENINLVE-ESGKP 422

394 KE-CLNQTSKAVHNDVSYSNIKGTIVRTPIHFACSDTVACTNITLSEVELLPFGALL 452

423 SEATCKNVHFNNAHYTPHCTSLSEDEAL 453

453 DDFCMNAVGTQRTLTIPINCAREGDPETV 483

Search completed: March 24, 2005, 07:36:03

Job time: 1258 sec







```

Db      361 YCDRVEPCIOQFSANQVKNVYENIKGTSATKAIKFDGCTNPGGIMENINLVGSG 420
Qy      421 KPSEATCKNVHFNNAEHVTPHCTSLSEDEALLVNY 457
Db      421 KPSEATCKNVHFNNAEHVTPHCTSLSEDEALLVNY 457

```

## RESULT 2

```

5447867-1
; Patent No. 5447867
; APPLICANT: BRIDGES, IAN; SCHUCH, WOLFGANG; GRIERSON, DONALD
; TITLE OF INVENTION: RECOMBINANT DNA CONTAINING PECTIN
; ESTERASE GENE SEGMENTS
; NUMBER OF SEQUENCES: 4
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/24,866
; FILING DATE: 26-FEB-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 720,629
; FILING DATE: 25-JUN-1991
; APPLICATION NUMBER: 419,779
; FILING DATE: 29-SEP-1989
; APPLICATION NUMBER: 119,614
; FILING DATE: 12-NOV-1987
; SEQ ID NO:1
; LENGTH: 457
5447867-1

```

```

Query Match      100.0%; Score 2390; DB 6; Length 457;
Best Local Similarity 100.0%; Pred. No. 8,6e-208;
Matches 457; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 MYIQRNSILLIIIFASSISTCRSNVIDNLFPQVYDNLLEQEPHADFQAVLYSKNIE 60
Db      1 MYIQRNSILLIIIFASSISTCRSNVIDNLFPQVYDNLLEQEPHADFQAVLYSKNIE 60
Qy      61 SNNNIDKVDKNGIKVYNLUSFGAKGDKTYDNIAEQAMNEACSRTPVQFVVPKXNYL 120
Db      61 SNNNIDKVDKNGIKVYNLUSFGAKGDKTYDNIAEQAMNEACSRTPVQFVVPKXNYL 120
Qy      121 LKQITTSFGPCRSSISIVKIFGSLSEASSKISDYKDRMLTAEPSONLVYGGGTTNGGOY 180
Db      121 LKQITTSFGPCRSSISIVKIFGSLSEASSKISDYKDRMLTAEPSONLVYGGGTTNGGOY 180
Qy      181 WMPSSCKIKKSLPCGDAPTALTLPNCKNLKVNNLKSNQAQIHIKFESECTNVVASMLMIN 240
Db      181 WMPSSCKIKKSLPCGDAPTALTLPNCKNLKVNNLKSNQAQIHIKFESECTNVVASMLMIN 240
Qy      241 ASAKSPNTDGVHVSNTQYIQISDTIIGTDDCISIVSGSONVQATNITGPGHGISIGSL 300
Db      241 ASAKSPNTDGVHVSNTQYIQISDTIIGTDDCISIVSGSONVQATNITGPGHGISIGSL 300
Qy      301 GSGNSEAVNVNTEAKIIGAENGVRITKWOSSGQASNIKFLNVEMQDYKPIIIDON 360
Db      301 GSGNSEAVNVNTEAKIIGAENGVRITKWOSSGQASNIKFLNVEMQDYKPIIIDON 360
Qy      361 YCDRVEPCIOQFSANQVKNVYENIKGTSATKAIKFDGCTNPGGIMENINLVGSG 420
Db      361 YCDRVEPCIOQFSANQVKNVYENIKGTSATKAIKFDGCTNPGGIMENINLVGSG 420
Qy      421 KPSEATCKNVHFNNAEHVTPHCTSLSEDEALLVNY 457
Db      421 KPSEATCKNVHFNNAEHVTPHCTSLSEDEALLVNY 457

```

## RESULT 3

```

US-09-051-239A-2
; Sequence 2, Application US/09051239A
; Patent No. 6420628
; GENERAL INFORMATION:
; APPLICANT: ULVSKOV, Peter
; APPLICANT: CHILD, Robin

```

```

; APPLICANT: VAN ONCKELIN, Henri
; APPLICANT: PRINSEN, Els
; APPLICANT: BORKHARDT, Bernard
; APPLICANT: SANDER, Lilli
; APPLICANT: PETERSEN, Merten
; APPLICANT: BUNDGARD POUlsen, Gert
; APPLICANT: BOTTERMAN, Johan
; TITLE OF INVENTION: Seed Shattering
; FILE REFERENCE: 2121-0138P
; CURRENT APPLICATION NUMBER: US/09/051,239A
; PRIOR FILING DATE: 1998-09-28
; PRIOR APPLICATION NUMBER: PCT/EP96/04313
; PRIOR FILING DATE: 1996-10-04
; PRIOR FILING DATE: 1995-10-06
; PRIOR APPLICATION NUMBER: EP 95 402241.4
; PRIOR FILING DATE: 1995-12-08
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 2
; LENGTH: 433
; TYPE: PRT
; ORGANISM: Brassica napus
; FEATURE:
; OTHER INFORMATION: Strain cv. Topaz.
US-09-051-239A-2

```

```

Query Match      46.6%; Score 1113; DB 4; Length 433;
Best Local Similarity 49.9%; Pred. No. 3,1e-92;
Matches 223; Conservative 69; Mismatches 121; Indels 34; Gaps 8;

Qy      7 SLILIIIFASSISTCRSNVIDNLFPQVYDNLLEQEPHADFQAVLYSKNIESNNID 66
Db      9 AVFLCVLMLALCCQALSSNV-DDG-----YGHEDG---SFEPSDLIRKLNDD 51
Qy      67 KY-----DKNGIKVYNLUSFGAKGDKTYDNIAEQAMNEACSRTPVQFVVPKXNYL 118
Db      52 DVLLKSDPTPTSESTSVSNFGAKGDKTIDTQAFKAWKKA CSTNGVTFILPKCKT 111
Qy      119 YLLKQITTSFGPCRSSISIVKIFGSLSEASSKISDY-KDRMLTAEPSONLVYGGG--GTIN 175
Db      112 YLLKSIIRFRGPKCKLSFQILGTLASSTRSDYSNDKXHWLILDDVNNLSIDGSGAGIVD 171
Qy      176 GNGQVWMPSSCKIKKSLPCGDAPTALTLPNCKNLKVNNLKSNQAQIHIKFESECTNVVAS 235
Db      172 GNGNIMWQNSCKIDSKPCTTAPALTLYNKNLNVKXLRVRNAQIQISIEKNNVGVK 231
Qy      236 NMTNINSAKSPNTDGVHVSNTQYIQISDTIIGTDDCISIVSGSONVQATNITGPGHG1 295
Db      232 NVKITAPOGSPNTDGIHIVATKNIRISNSDICTGDDCISIEDGSONVQINDLTGPGHG1 291
Qy      296 SIGSLGSGNSEAVNVNTEAKIIGAENGVRITKWOSSGQASNIKFLNVEMQDYKPI 355
Db      292 SIGSLGDDNSKAYVSGIVDADLTSETDNGVRITKYOGSGGTAKNIKFNIRMDVKNPI 351
Qy      356 IIDQVYCNVRPCTIQFSANQVKNVYENIKGTSATKAIKFDGCTNPGGIMENINLV 415
Db      352 IIDQVYCNVR-DKCQOQFSANQVKNVYENIKGTSATKAIKFDGCTNPGGIMENINLV 410
Qy      416 VGESGKPSEATCKNVHFNNAEHVTPHCT 442
Db      411 KG-----GKASCENNVVXKDKGVSPKC 432

```

## RESULT 4

```

US-10-151-668-2
; Sequence 2, Application US/10151668
; Patent No. 6797861
; GENERAL INFORMATION:
; APPLICANT: ULVSKOV, Peter
; APPLICANT: CHILD, Robin
; APPLICANT: VAN ONCKELIN, Henri
; APPLICANT: PRINSEN, Els

```

```

/ APPLICANT: BORKHARDT, Bernard
/ APPLICANT: SANDER, Lilli
/ APPLICANT: PETERSEN, Morten
/ APPLICANT: BUNDGARD, Poulsen, Gert
/ APPLICANT: BOTTERMANN, Johan
/ TITLE OF INVENTION: Seed Shattering
/ FILE REFERENCE: 2121-0138P
/ CURRENT APPLICATION NUMBER: US/10/151,668
/ PRIOR FILING DATE: 2002-05-21
/ PRIOR APPLICATION NUMBER: US/09/051,239
/ PRIOR FILING DATE: 1998-09-28
/ PRIOR APPLICATION NUMBER: PCT/EP96/04313
/ PRIOR FILING DATE: 1996-10-04
/ PRIOR APPLICATION NUMBER: EP 95 402241.4
/ PRIOR FILING DATE: 1995-10-06
/ PRIOR APPLICATION NUMBER: EP 95 203328.0
/ PRIOR FILING DATE: 1995-12-08
/ NUMBER OF SEQ ID NOS: 14
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 2
/ LENGTH: 433
/ TYPE: PRT
/ ORGANISM: Brassica napus
/ FEATURE:
/ OTHER INFORMATION: Strain cv. Topaz.
US-10-151-668-2

```

Query Match 46.6%; Score 1113; DB 4; Length 433;

Best Local Similarity 49.9%; Pred. No. 3.1e-92; Matches 223; Conservative 69; Mismatches 121; Indels 34; Gaps 8;

```

QY 7 STILLIIFASSISTCRSNVIDNLFKQVYDNILEQFAHDFQAYLSYLSKNIESNNID 66
DB 9 AIFLCVLMALACCOALSSNV-DG-----YGHEDG---SFEHDSILKLNND 51
QY 67 KV-----DKNGIKVINLVSPGAKDGKTYDNIAFEQAMNACSSRTVPQFVPRKNX 118
DB 52 DVLTKSDRPTRESSIVSNFGAKDGDGKTDDTQAFKKAMKACSTNGVTTFLLPKGKT 111
QY 119 YLLKQTFSPGCRSSISVKTFGSLBASKSIDY-KDRRLMIAFDSVQNLVVGSG--GTIN 175
DB 112 YLLKSIRFRGPKSLRSFQILGLTSLASTKRSYSDNKNHMLLEEDVNNLSIDGSGAGIYD 171
QY 176 GNGQWMPSSCKINKSLPCRDAPALTLPNNCKNLKVNLSKXNAQOIHKFESCTNVVAS 235
DB 172 GNNINWQNSCKIDSKSPCTKAPALTLYNLKLNKLVNNAQOQIISIEKCNVGVK 231
QY 236 NLMINASAKSPNTDGVHVSNTQYIQISDTIIGTGDDCISIVSGSONVQATNITCGPHGI 295
DB 232 NVKITAPGDSPTNDGHIHIVATKNIRISNSDIGTDDCISIEDGSONVQINDLTCGPHGI 291
QY 296 SIGSLGSGNSEAVSVNTVNEAKIIGAENGVRITKMOGSGGASNTKFLNVEHQDYKYP1 355
DB 292 SIGSLGSDNSKAVYSGIDVAGATLSEFTDNGVRITKYQSGGTAKNIKFQIRMDNVKNP1 351
QY 356 IIDQWCDRREPCTIOQFSAVQVKNVYENIKGTSATVAIKFDCSTNPFCEGIMENINL 415
DB 352 IIDQWCDK-DKCEQGESAVQVNNVYRNIGTSATDAIVMFCSSVKYPCQGLVLENNV1 410
QY 416 VGESGKPSKATCNVFNNAEHTPHC 442
DB 411 KG-----GKASCKNVNVKDKGTIVSPKC 432

```

```

REGUL 5
US-08-941-532-6
/ Sequence 6, Application US/08941532
/ Patent No. 6096946
/ GENERAL INFORMATION:
/ APPLICANT: ROBERTS, Jeremy Alan
/ APPLICANT: COUPE, Simon Allan
/ APPLICANT: JENKINS, Elizabeth Sarah
/ TITLE OF INVENTION: CONTROL OF POD DEHISCENCE

```

```

/ NUMBER OF SEQUENCES: 8
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Sterne, Kessler, Goldstein & Fox
/ STREET: 1100 New York Avenue
/ CITY: Washington
/ STATE: D.C.
/ COUNTRY: U.S.A.
/ ZIP: 20005
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/941,532
/ FILING DATE: 30-SEP-1997
/ CLASSIFICATION: 800
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: PCT/GB96/00757
/ FILING DATE: 29-MAR-1996
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: GB 9506684.1
/ FILING DATE: 31-MAR-1995
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Esmond, Robert W.
/ REGISTRATION NUMBER: 32,893
/ REFERENCE/DOCKET NUMBER: 0623.0580001/RWE
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (202) 371-2600
/ TELEFAX: (202) 371-2540
/ INFORMATION FOR SEQ ID NO: 6:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 433 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
US-08-941-532-6

```

Query Match 46.2%; Score 1103; DB 3; Length 433;

Best Local Similarity 49.8%; Pred. No. 2.5e-91; Matches 221; Conservative 70; Mismatches 125; Indels 28; Gaps 7;

```

QY 7 STILLIIFASSISTCRSNVID-----DNLFKQVYDNILEQFAHDFQAYLSYLSKNIES 61
DB 9 AIFLCVLMALACCOALSSNVDDYGHEDGSEF--DSLTKLNNDVLLTKSSDRPTES 66
QY 62 NNNIDKVRDKNGIKVINLVSPGAKDGKTYDNIAFEQAMNACSSRTVPQFVPRKNXLYL 121
DB 67 S-----TVSVSNFGAKDGDGKTDDTQAFKKAMKACSTNGVTTFLLPKGTYLL 114
QY 122 KQTFSPGCRSSISVKTFGSLBASKSIDY-KDRRLMIAFDSVQNLVVGSG--GTINGN 178
DB 115 KSTIRFGPKSLRSFQILGLTSLASTRSDYSDNKNHMLLEEDVNNLSIDGSGAGIYDNG 174
QY 179 QVWMPSSCKINKSLPCRDAPALTLPNNCKNLKVNLSKXNAQOIHKFESCTNVVASNLM 238
DB 175 KIMWQNSCKIDSKSPCTKAPALTLYNLKLNKLVNNAQOQIISIEKCNVGVK 234
QY 239 INASAKSPNTDGVHVSNTQYIQISDTIIGTGDDCISIVSGSONVQATNITCGPHGISIG 298
DB 235 ITAPGDSPTNDGHIHIVATKNIRISNSDIGTDDCISIEDGSONVQINDLTCGPHGISIG 294
QY 299 SIGSGNSEAVSVNTVNEAKIIGAENGVRITKMOGSGGASNTKFLNVEHQDYKYP1IID 358
DB 295 SIGDDNSKAVYSGIDVAGATLSEFTDNGVRITKYQSGGTAKNIKFQIRMDNVKNP1IID 354
QY 359 QNYCDRREPCTIOQFSAVQVKNVYENIKGTSATVAIKFDCSTNPFCEGIMENINLVGE 418
DB 355 QNYCDK-DKCEQGESAVQVNNVYRNIGTSATDAIVMFCSSVKYPCQGLVLENNV1KG- 412
QY 419 SGKPSKATCNVFNNAEHTPHC 442
DB 413 ---GKASCKNVNVKDKGTIVSPKC 432

```

RESULT 6  
US-08-467-023-134  
; Sequence 134, Application US/08467023  
; Patent No. 6090386  
; GENERAL INFORMATION:  
; APPLICANT: Griffee, Irwin J.;  
; APPLICANT: Pollock, Joanne;  
; APPLICANT: Bond, Julian F.;  
; APPLICANT: Garman, Richard D;  
; APPLICANT: Kuo, Mei-Chang;  
; APPLICANT: Yeung, Siu-mei H.;  
; APPLICANT: Brauer, Andrew;  
; APPLICANT: Exley, Mark A.;  
; TITLE OF INVENTION: Allergenic Proteins And Peptides From  
; TITLE OF INVENTION: Japanese Cedar Pollen  
; NUMBER OF SEQUENCES: 261  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Immunologic Pharmaceutical Corporation, Inc.  
; STREET: 610 Lincoln St  
; CITY: Waltham  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02154  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/467,023  
; FILING DATE: June 6, 1995  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/350,225  
; FILING DATE: December 6, 1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Jane E. Remillard  
; REGISTRATION NUMBER: 38,872  
; REFERENCE/DOCKET NUMBER: 025.6 USD2 (IM1-028CPD2)  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 227-5941  
; TELEFAX: (617) 227-5941  
; INFORMATION FOR SEQ ID NO: 134:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 514 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-467-023-134

Query Match 32.9%; Score 787; DB 3; Length 514;  
Best Local Similarity 40.5%; Pred. No. 1,3e-62;  
Matches 168; Conservative 73; Mismatches 160; Indels 14; Gaps 7;

QY 45 AHPQAVL---SYLSNINNNNDVYD---KNGIKVINVLSEFGAKGDKGTVDNIAFEQA 98  
DB 22 AEDQSAQIMLDSVVEKYLRSNRSIRKVEHSRHDANINPVEKYGAVGKGDCTEAFSTA 81  
QY 99 KVNACSSRTFVQVYKKNKNYLLKQITTSGPCRSSISVYKFGSLSEASSKSIDYKDRRLMI 158  
DB 82 WQACACNBS-ANMLVPGSKKFFVNNLFFVPGPCOPHTEFKVDGIIAAYQNPAWSKNNRIWL 140  
QY 159 AFDVQNTLVVGGGTTINGNGVWMPSSCK-INKSLPC--RAPATLTLTWNCKATKVNKLK 215  
DB 141 QFAKLTGFTLMGKVIDGQGXWAGCKWVNGREICNDRDPATKFDSTGLIIOGLK 200  
QY 216 SQAAGQIHFKESCTNVVAVSNLMINASAKSPNTGVAHSNTQYIOISPTLITGDCDCISI 275  
DB 201 LMSBPEFHLVFQNGCGVKITIGISITAPRDSPTDGDIDIFASKNPHLQKNTIGTDDCVAI 260

QY 276 VSGSNVQATNTITGPGHGISTIGSGNSEAYVNTVNEAKIIGAENGRIKTNQGS 335  
DB 261 GTGSSNIVIEDLIGPBGHGISISLNGENSRARVSYHVNGAKFIDTONGLRITKTNQGS 320  
QY 336 GQASNIKFLVEMQDVKYPITIDONYCDVRPECTIOQSAVQVNVYENIKGTSATKVAI 395  
DB 321 GMASHIYEVNEMINSENPIITINQFCTASASACQNRSAVQIDYVYKIRGTSATPAAI 380  
QY 396 KFDCTNTPQEGIMENINVGESGKPSKATCNVFN--AEHTPHCTSLKIS 448  
DB 381 QKQSDSMPCKDKISLISIKLTSGK--IASCLINDNANGVFSGHVLPACKNLSPS 433

RESULT 7  
5453566-2  
; Patent No. 5453566  
; APPLICANT: SHENMAKER, CHRISTINE K.; KRIDL, JEAN C.; HIATT,  
; WILLIAM R.; KNAUF, VIC  
; TITLE OF INVENTION: ANTISENSE REGULATION OF GENE EXPRESSION  
; IN PLANT/CELLS  
; NUMBER OF SEQUENCES: 2  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/750,505  
; FILING DATE: 27-AUG-1991  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 240,408  
; FILING DATE: 30-AUG-1988  
; APPLICATION NUMBER: 920,574  
; FILING DATE: 17-OCT-1986  
; APPLICATION NUMBER: 845,676  
; FILING DATE: 28-MAR-1986  
; SEQ ID NO: 2  
; LENGTH: 70

Query Match 15.3%; Score 365; DB 6; Length 70;  
Best Local Similarity 100.0%; Pred. No. 1.1e-25;  
Matches 70; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 72 GIKYINVLSPAKGDKTYDNIATFEQAMNACSSRTPVQVVPKKNKNYLLKQITTSFGPCR 131  
DB 1 GIKYINVLSPAKGDKTYDNIATFEQAMNACSSRTPVQVVPKKNKNYLLKQITTSFGPCR 60  
QY 132 SSISSVIFGS 141  
DB 61 SSISSVIFGS 70

RESULT 8  
5453566-2  
; Patent No. 5453566  
; APPLICANT: SHENMAKER, CHRISTINE K.; KRIDL, JEAN C.; HIATT,  
; WILLIAM R.; KNAUF, VIC  
; TITLE OF INVENTION: ANTISENSE REGULATION OF GENE EXPRESSION  
; IN PLANT/CELLS  
; NUMBER OF SEQUENCES: 2  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/750,505  
; FILING DATE: 27-AUG-1991  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 240,408  
; FILING DATE: 30-AUG-1988  
; APPLICATION NUMBER: 920,574  
; FILING DATE: 17-OCT-1986  
; APPLICATION NUMBER: 845,676  
; FILING DATE: 28-MAR-1986  
; SEQ ID NO: 2  
; LENGTH: 70

Query Match 15.3%; Score 365; DB 6; Length 70;  
Best Local Similarity 100.0%; Pred. No. 1.1e-25;  
Matches 70; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 72 GIKIVNLSFGAKGDKTYDNIAPFOAMNEACSSRTPVQVVPKKNYLLKQITFSGPCR 131  
DB 1 GIKIVNLSFGAKGDKTYDNIAPFOAMNEACSSRTPVQVVPKKNYLLKQITFSGPCR 60  
QY 132 SSISVKIFGS 141  
DB 61 SSISVKIFGS 70

RESULT 9  
US-08-467-023-189  
Sequence 189, Application US/08467023  
Patent No. 6090386  
GENERAL INFORMATION:  
APPLICANT: Griffee, Irwin J.;  
APPLICANT: Pollock, Joanne;  
APPLICANT: Bond, Julian F.;  
APPLICANT: Garman, Richard D;  
APPLICANT: Kuo, Mei-Chang;  
APPLICANT: Yeung, Shu-mei H.;  
APPLICANT: Brauer, Andrew;  
APPLICANT: Exley, Mark A.;  
APPLICANT: Powels, Steven P.  
TITLE OF INVENTION: Allergenic Proteins And Peptides From  
TITLE OF INVENTION: Japanese Cedar Pollen  
NUMBER OF SEQUENCES: 261  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Immunologic Pharmaceutical Corporation, Inc.  
STREET: 610 Lincoln St  
CITY: Waltham  
STATE: MA  
COUNTRY: USA  
ZIP: 02154  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/467,023  
FILING DATE: June 6, 1995  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/350,225  
FILING DATE: December 6, 1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Jane E. Remillard  
REGISTRATION NUMBER: 38,872  
REFERENCE/DOCKET NUMBER: 025.6 USD2 (IMI-028CPD2)  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 227-7400  
TELEFAX: (617) 227-5941  
INFORMATION FOR SEQ ID NO: 189:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 127 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FRAGMENT TYPE: internal  
US-08-467-023-189

Query Match 13.3%; Score 319; DB 3; Length 127;  
Best Local Similarity 56.0%; Pred. No. 3.8e-21;  
Matches 65; Conservative 16; Mismatches 35; Indels 0; Gaps 0;

QY 293 HGISISLGSNEAVSYNTVNEAKIIGANGVRIRIKTQSGSGAANIETFLVEMQDYK 352  
DB 8 HESISISLGSNEAVSYNTVNEAKIIGANGVRIRIKTQSGSGAANIETFLVEMQDYK 67  
QY 353 YPIIDONTCDRVEPCIQQFSAAVQKVVYENIKGTISATKVAIKFDCSTNPFCEGI 408  
DB 68 NPILINQPCYTSASACONQBSAVQIDVYTKNIRGTSATPAAIQLKCSDSMPCKDI 123

RESULT 10  
US-09-107-532A-6652  
Sequence 6652, Application US/09107532A  
Patent No. 6563275  
GENERAL INFORMATION:  
APPLICANT: Lynn A Doucette-Stamm and David Bush  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO  
ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS  
NUMBER OF SEQUENCES: 7310  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: GENOME THERAPEUTICS CORPORATION  
STREET: 100 Beaver Street  
CITY: Waltham  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02354  
COMPUTER READABLE FORM:  
MEDIUM TYPE: CD-ROM ISO9660  
COMPUTER: PC  
OPERATING SYSTEM: <Unknown>  
SOFTWARE: ASCII  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/107,532A  
FILING DATE: 30-Jun-1998  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/085,598  
FILING DATE: 14 May 1998  
APPLICATION NUMBER: 60/051571  
FILING DATE: July 2, 1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Ariniello, Pamela Deneke  
REGISTRATION NUMBER: 40,489  
REFERENCE/DOCKET NUMBER: GTC-012  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (781)893-5007  
TELEFAX: (781)893-8277  
INFORMATION FOR SEQ ID NO: 6652:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 451 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: YES  
ORIGINAL SOURCE:  
ORGANISM: Enterococcus faecium  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (B) LOCATION 1...451  
SEQUENCE DESCRIPTION: SEQ ID NO: 6652:  
US-09-107-532A-6652

Query Match 12.5%; Score 299; DB 4; Length 451;  
Best Local Similarity 27.2%; Pred. No. 1.6e-18;  
Matches 99; Conservative 65; Mismatches 132; Indels 68; Gaps 15;

QY 73 IXYINLSFGAKGDKTYDNIAPFOAMNEACSSRTPVQVVPKKNYLLKQITFSGPCR 132  
DB 5 IMYDILTFGASID--ELNTEAIOQAIDMAASDGGT--VVVPGE-----FLTGALFL 54  
QY 133 SISVKIFGSLBASKIS--DYKD-----RLW---IAFDSVONLVVGGGTING 176  
DB 55 KSNVEIHLASGAVLKSDDPKDYPPVHSHRVEGHRKRVASCIYAQVENVISVGFETLDS 114  
QY 177 NGQVW-----PSGCKINSLPCRDAPYALTFWNCNKLKVNLSKNAQOIHIKESCT 230  
DB 115 NGKQWHTFRNEPD---NLAYP---RPLKSPHNRCHRTIVDXIKLIGSPSWTINPLCS 167  
QY 231 NVVASNLMNAGAKSNTGVAVSNTQYIQISDTIIGTDDDCISYSGS-----QN 281  
DB 168 NATFDNLTLINPADSNTDGIPESCKNVRISNCHDIVDDCIAIAGTEDIYRIACEN 227





Best Local Similarity 25.1%; Pred. No. 1.8e-18;  
Matches 107; Conservative 73; Mismatches 166; Indels 61; Gaps 16;

QY 74 KYINVLSPFAGKDGKTYNDIAFEQANNEACSSRTPOVFPVKKNKYL--KOITSSGC 130  
DB 48 KCHVASHSDGTDSDYILSLNQ-----CNHGKRVF--DEBKEYITGLMLMTF---- 96  
QY 131 RRSISVKINGSL-----EASSKISDYKDRRLMIAPDSVONLVVGGGTTINGGOYW 181  
DB 97 LKNIDLEVLTGTLFTDTDTYQANNSFKQGFQNFATTFQJLGG-EDVMWGGJITINGGOYW 155  
QY 182 WPSSCKINKSLPCRDAPFALTFFWNCNKLKYNLKSNAQOIHIFESCTNVVASNLMINA 241  
DB 156 YDLVAEDDIL-----RPLMGIILGANGGTIGLKLKVPYHFVANSVLFDGIDISG 211  
QY 242 SAKS---PNTDGVHVSNTQYIOISDTITGDDPCISYSGSQNVQATNTGCPHGISI 297  
DB 212 YSKSDNEAKTKDGMWDTYRSNNIVIONSVINNGDCVSPKPNSTNILLVQNLCHGSHGISV 271  
QY 298 GSLGSGNSE-----AYVSNVTNNAKIIIGAENGV-----RIKTNQ----- 332  
DB 272 GSLGQYKDEVDIYENVYVNIISMFNASVCLNFHHIIDPLTLQDMARIKVMFGTSSALS 331  
QY 333 -----GSSGOASNIKFLNVMQDVYPIIIDQNYCDR-VEPCIQFSAVOVKVYVENI 385  
DB 332 ADLQGGGGSGSVKNITVDALINVDMAIEITQCYGQKNTLGNVPSSTISDVHIKMF 391  
QY 386 KGTSA-----TKVAKFDCSTNFPCEGIIMENINLVGSGKSEBATCKVHFNNAAH 437  
DB 392 RGTTSSEDPYVGTIV-----CSSPDTCSDIYNSINIVSPDG-TNDFVCBNV---DESL 442  
QY 438 VTFPHCTS 444  
DB 443 LSVNCTA 449

RESULT 13  
US-09-787-583-2  
Sequence 2, Application US/09787583  
Patent No. 6602696  
GENERAL INFORMATION:  
APPLICANT: DSM N.V.  
APPLICANT: Franse, Maartje  
APPLICANT: Graessin, Catherine  
APPLICANT: Herweijer, Margareta  
APPLICANT: Meelwisen, Petrus  
APPLICANT: Ooljen, Albert  
APPLICANT: Voragen, Alphons  
TITLE OF INVENTION: ASPERGILLUS TUBIGENSIS POLYGALACTURONASE  
FILE REFERENCE: 24615-20144.00  
CURRENT APPLICATION NUMBER: US/09/787,583  
PRIOR FILING DATE: 2001-03-19  
PRIOR APPLICATION NUMBER: EP 99200481.2  
PRIOR FILING DATE: 1999-02-17  
PRIOR APPLICATION NUMBER: EP 98203171.8  
PRIOR FILING DATE: 1998-09-18  
NUMBER OF SEQ ID NOS: 2  
SOFTWARE: FaalSeq for Windows Version 4.0  
SEQ ID NO 2  
LENGTH: 362  
TYPE: PRT  
ORGANISM: Aspergillus tubingensis  
US-09-787-583-2

Query Match 12.3%; Score 293.5; DB 4; Length 362;  
Best Local Similarity 28.3%; Pred. No. 3.6e-18;  
Matches 80; Conservative 59; Mismatches 91; Indels 53; Gaps 13;

QY 171 GGTINNGOVW-----PSSCKINKSLPCRDAPFALTFFWNCNKLKYNLKSNA 219  
DB 102 GAKINCDGARWMDGKSGNGKTRPKFFQVHK-----LDSSITGKIKYNTPTVG- 150  
QY 220 QQHIKIFESCTNVVASNLMIN-----ASAKSPMTDGVHVSNTQYIOISDTITGDD 271

DB 151 -----FSLADHLTITDVTIDNSAGTSKGNHTDAFDIGQSTYITIDGATVNOOD 200  
QY 272 CTSYSGSQNVQATNTGCPHGICISIGLSGNSSEAYVSNVTVNNAKIIIGAENGVRKTM 331  
DB 201 CLAINSG-EHITFTNGYCGGHLGSLG-GRSDVTVDVITISNKLNSQNGVRIKTI 258  
QY 332 OGSQGOASNIKFLNVMQDV-KYPIIIDQNYCDRVEPCIQFSAVOVKVYVENIKG--- 387  
DB 259 YKRTGVEVNEKFDITLSDISKGIIVEODY-ENGSPDTGPTINGVVEEDITFKKVTGSVK 317  
QY 388 TSATKVAIKFDCSTNFPCEGIIMENINLVGSGKSEATCKV 430  
DB 318 SSGTDIYIL--CGSG-SCSNMTSGVDYTG--GKKS-SKCKV 354

RESULT 14  
US-09-542-767A-1  
Sequence 1, Application US/09542767A  
Patent No. 6296671  
GENERAL INFORMATION:  
APPLICANT: Schuelein, Martin  
APPLICANT: Kriebensen, Henrik  
TITLE OF INVENTION: An Enzymatic Treatment Method  
FILE REFERENCE: 5871.204-US  
CURRENT APPLICATION NUMBER: US/09/542,767A  
PRIOR FILING DATE: 2000-04-04  
PRIOR APPLICATION NUMBER: PA 1999 00390  
PRIOR FILING DATE: 1999-03-22  
PRIOR APPLICATION NUMBER: US 60/125,884  
PRIOR FILING DATE: 1999-03-24  
PRIOR APPLICATION NUMBER: PCT/DK00/00136  
PRIOR FILING DATE: 2000-03-22  
NUMBER OF SEQ ID NOS: 1  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 1  
LENGTH: 360  
TYPE: PRT  
ORGANISM: Trichisporon penicillatum  
US-09-542-767A-1

Query Match 10.8%; Score 257.5; DB 3; Length 360;  
Best Local Similarity 25.5%; Pred. No. 6.5e-15;  
Matches 72; Conservative 54; Mismatches 105; Indels 51; Gaps 10;

QY 171 GGTINNGOVWPPSSCKINKSL-PCRDAPFALTFFWNCNKLKYNLKSNAQOIHIFESC 229  
DB 104 GSVTDEGARWD-----NKGANGGVKPKRLFYAHNLDSHINGLHIKXTPVFGSIDS- 157  
QY 230 TTVVASNLMINASAKSP-----NTDGVHVSNTQYIOISDTITGDDPCISYSGSQNVQAT 285  
DB 158 KNLIDGVARIDNSDGTQCAFNTDADVQSQVYVTTIQNMAWNOODCLAINQ-ELIHFL 216  
QY 286 NITCGPHGISIGLSGNSSEAYVSNVTVNNAKIIIGAENGVRKTMQSGSQASNIKFLN 345  
DB 217 NGCYGCGHLSIGSVGGN---VSDVVIADSQINSQGVRIKTRSGCGVEVRIYRN 273  
QY 346 VEMQDV-KYPIIIDQNYCDRVEPCIQFSAVOVKVYVENIKTSA----- 390  
DB 274 IFLSGITDYGLIVOODYNNPQHAT---NSIKHIDITFPNVHGTATQHGFNIAIFCGDGS 329  
QY 391 ----TKVAKFDCSTNFPCEGIIMENINLVGSGKSEATCK 428  
DB 330 CIDWTWNEVKIHGARDYCKQNV-----PSSASCQ 358

RESULT 15  
US-08-061-062A-6  
Sequence 6, Application US/08061062A  
Patent No. 5550045  
GENERAL INFORMATION:  
APPLICANT: MUSTERS, WOUTER  
APPLICANT: STAM, HEIN

us-10-691-374-2.rai

```

1  APPLICANT: SUYKERBUIK, MARIA E.
2  APPLICANT: VISSER, JACOB
3  APPLICANT: VERBAKEL, Johannes M.
4  TITLE OF INVENTION: CLONING AND EXPRESSION OF DNA
5  TITLE OF INVENTION: ENCODING A RIPENING FORM OF A POLYPEPTIDE HAVING
6  TITLE OF INVENTION: RHMNOALACTURONASE ACTIVITY
7  NUMBER OF SEQUENCES: 16
8  CORRESPONDENCE ADDRESS:
9  ADDRESSEE: CUSHMAN DARBY & CUSHMAN
10 STREET: 1100 NEW YORK AVENUE, N.W.
11 CITY: WASHINGTON, D.C.
12 COUNTRY: U.S.A.
13 ZIP: 20005-3918
14
15 COMPUTER READABLE FORM:
16 MEDIUM TYPE: Disk
17
18 COMPUTER: IBM PC compatible
19 OPERATING SYSTEM: PC-DOS/MS-DOS
20 SOFTWARE: PatentIn Release #1.0, Version #1.25
21 CURRENT APPLICATION DATA:
22 APPLICATION NUMBER: US/08/061,062A
23 FILING DATE: 14 MAY 1993
24 CLASSIFICATION: 435
25 ATTORNEY/AGENT INFORMATION:
26 NAME: KOKULIS, PAUL N.
27 REGISTRATION NUMBER: 16773
28 REFERENCE/DOCKET NUMBER: 202390/R 7262 (V)
29 TELECOMMUNICATION INFORMATION:
30 TELEPHONE: (202) 861-3000
31 TELEFAX: (202) 822-0944
32 TELEX: 6714627 CUSH
33 INFORMATION FOR SEQ ID NO: 6:
34 SEQUENCE CHARACTERISTICS:
35 LENGTH: 440 amino acids
36 TYPE: amino acid
37 TOPOLOGY: linear
38
39 MOLECULE TYPE: protein
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
95
```

QY 70 KNGIKVNVLSFGAKDDGKTYDNIATFEQAMNEACSSRTVPQVPEVPKNNVILKQITFSG 129  
 Db 33 KAKTKCNILSYGAVADNSTDVGPAITSAW-AAKCSGGLV-YIISGVYALNTWVTLTGG 89  
 QY 130 CRRSISVK--IFGSLERASSKISDYCKDRRLMIFDVSQVLVVGGGGTINGNQVWPSPSCK 187  
 Db 90 SATAIQLDGIIVRTGPAASGNMTAVTDTDFELFSSITSK-----GAQVGFYVYHABGT 143  
 QY 188 INKSLPCRPAPALIFWNCNKLKVNLLSKNAOQIHIFESCTVNVVASNIMTANSAKSPN 247  
 Db 144 GARILRLDV-----THRSVADVILMDAPAFHPTMDCISGEVYVNAIKRGNEG-G 193  
 QY 248 TDGVHVSNTVOYLOISDTLIIGTGDDCISIVSSQVNVQATNITGCGHGHSIGSLGSGNSEA 307  
 Db 194 LDGIDVWGSN-IWHDDEVNTKDECVTVKSPANNILVESIYCWMVSGGCAWGLG---ADT 249  
 QY 308 YVSNVTVNEAKIIQAENGVRITKTYQWGSQGAASNIKFLANVEMQVKKPIIIDQNYCDRVEP 367  
 Db 250 DVTDIYVRNNTVYSSNOMYIKS-NGSGSYVSNVLENTFGHNAVSLDIDGWSMTAV 308  
 QY 368 CIOQFSAVQVKNVYENIKGT-----SATKAIAKDCSTNFPCEGIIMENINLVGSGKPSB 424  
 Db 309 A---GDGEVLNNTIVKMKGTGANATRPRIRVVCSDTAPCTDLTLEDAIATWESGSSBL 365  
 QY 425 ATCKNVH 431  
 Db 366 YLCRSAY 372

Search completed: March 24, 2005, 07:06:49  
Job time : 45 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 24, 2005, 06:59:04 ; Search time 169 Seconds  
(without alignments)  
1045.855 Million cell updates/sec

Title: US-10-691-374-2

Perfect score: 2390

Sequence: 1 MWIGNSILLILIIIFASIS.....VPHCTSLSEIDELVNY 457

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: A\_Geneseq\_16Dec04:\*
- 2: geneseqp1980s:\*
- 3: geneseqp1990s:\*
- 4: geneseqp2000s:\*
- 5: geneseqp2001s:\*
- 6: geneseqp2002s:\*
- 7: geneseqp2003as:\*
- 8: geneseqp2003bs:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

| Result No. | Score  | Query Match | Length | DB ID | Description |
|------------|--------|-------------|--------|-------|-------------|
| 1          | 2390   | 100.0       | 457    | 1     | AAp80299    |
| 2          | 2390   | 100.0       | 457    | 1     | AAp94619    |
| 3          | 2390   | 100.0       | 457    | 2     | AAp32107    |
| 4          | 2390   | 100.0       | 457    | 4     | AAp84338    |
| 5          | 2338.5 | 97.8        | 456    | 6     | ABp43936    |
| 6          | 1185   | 49.6        | 438    | 5     | ABp92243    |
| 7          | 1127   | 47.2        | 431    | 2     | AAp98178    |
| 8          | 1127   | 47.2        | 431    | 3     | AAp41324    |
| 9          | 1127   | 47.2        | 431    | 5     | ABp92704    |
| 10         | 1119.5 | 46.8        | 415    | 3     | AAp41323    |
| 11         | 1119.5 | 46.8        | 415    | 3     | AAp41325    |
| 12         | 1115.5 | 46.2        | 426    | 5     | ABp92075    |
| 13         | 1103   | 46.2        | 433    | 2     | AAp04268    |
| 14         | 1103   | 46.2        | 433    | 2     | AAp42649    |
| 15         | 1056.5 | 44.2        | 404    | 2     | AAp48674    |
| 16         | 925    | 38.7        | 452    | 4     | AAp00422    |
| 17         | 888.5  | 37.2        | 441    | 5     | ABp91673    |
| 18         | 845    | 35.4        | 436    | 3     | AAp52347    |
| 19         | 845    | 35.4        | 436    | 3     | AAp52346    |
| 20         | 845    | 35.4        | 436    | 3     | ABp91108    |
| 21         | 840.5  | 35.2        | 436    | 5     | ABp91537    |
| 22         | 796.5  | 33.3        | 514    | 2     | AAp42122    |
| 23         | 791.5  | 32.9        | 514    | 2     | AAp04346    |
| 24         | 787.5  | 32.9        | 434    | 5     | ABp91377    |
| 25         | 787.5  | 32.9        | 514    | 2     | AAp74333    |

|    |       |      |     |   |          |          |           |
|----|-------|------|-----|---|----------|----------|-----------|
| 26 | 787.5 | 32.9 | 514 | 2 | AAp81586 | AAp81586 | Cedar pol |
| 27 | 787.5 | 32.9 | 514 | 2 | AAp25667 | AAp25667 | Japanese  |
| 28 | 787.5 | 32.9 | 514 | 7 | ADp34915 | ADp34915 | Cedar all |
| 29 | 787   | 32.9 | 514 | 2 | AAp53690 | AAp53690 | Japanese  |
| 30 | 787   | 32.9 | 514 | 2 | AAp69792 | AAp69792 | Japanese  |
| 31 | 787   | 32.9 | 514 | 2 | AAp93599 | AAp93599 | Japan ced |
| 32 | 787   | 32.9 | 514 | 2 | AAp25666 | AAp25666 | Japanese  |
| 33 | 787   | 32.9 | 514 | 7 | ADp34914 | ADp34914 | Cedar all |
| 34 | 767   | 32.1 | 460 | 2 | AAp69791 | AAp69791 | Japanese  |
| 35 | 764   | 32.0 | 453 | 4 | AAp51693 | AAp51693 | Japanese  |
| 36 | 764   | 32.0 | 507 | 4 | AAp51691 | AAp51691 | Japanese  |
| 37 | 741.5 | 31.0 | 491 | 5 | ABp90812 | ABp90812 | Herbicide |
| 38 | 728   | 30.5 | 405 | 5 | ABp92113 | ABp92113 | Herbicide |
| 39 | 724.5 | 30.3 | 398 | 5 | AAp20568 | AAp20568 | Pear poly |
| 40 | 718   | 30.0 | 435 | 5 | ABp93472 | ABp93472 | Herbicide |
| 41 | 718   | 30.0 | 482 | 3 | AAp05118 | AAp05118 | Arbidope  |
| 42 | 715   | 29.9 | 392 | 3 | ABp92116 | ABp92116 | Herbicide |
| 43 | 709.5 | 29.7 | 515 | 5 | ABp90945 | ABp90945 | Herbicide |
| 44 | 708.5 | 29.6 | 401 | 3 | AAp41452 | AAp41452 | Arbidope  |
| 45 | 708.5 | 29.6 | 422 | 3 | AAp41451 | AAp41451 | Arbidope  |

# ALIGNMENTS

|                       |  |
|-----------------------|--|
| RESULT 1              |  |
| AAp80299              |  |
| ID                    | AAp80299 standard; protein; 457 AA.  |
| XX                    |  |
| AC                    | AAp80299;  |
| XX                    |  |
| DT                    | 25-MAR-2003 (revised)  |
| DT                    | 20-NOV-1990 (first entry)  |
| XX                    |  |
| DE                    | Polygalacturonase.   |
| XX                    |  |
| KW                    | Fruit ripening; polygalacturonase; pectin esterase.  |
| XX                    |  |
| OS                    | Unidentified.  |
| XX                    |  |
| PN                    | EP271988-A.  |
| XX                    |  |
| PD                    | 22-JUN-1988.   |
| XX                    |  |
| PF                    | 06-NOV-1987; 87EP-00309853.  |
| XX                    |  |
| PR                    | 11-NOV-1986; 86GB-00026879.  |
| XX                    |  |
| PA                    | (ICIL ) IMPERIAL CHEM IND PLC.   |
| PA                    | (ZENE ) ZENCA LTD.   |
| PI                    | Bridges IG, Schuch WW, Grierson D;   |
| XX                    |  |
| DR                    | WPI: 1988-169271/25.   |
| DR                    | N-PSDB; AAN80487.  |
| XX                    |  |
| PT                    | Recombinant DNA comprising promoter and terminator sequences - useful in plants for altering ripening properties esp. in tomatoes.   |
| XX                    |  |
| PS                    | Disclosure; Page 7; 22pp; English.   |
| XX                    |  |
| CC                    | This polygalacturonase (PG) is encoded by plasmid clone pGOW6 which is used to produce antisense mRNA (with an inverted sequence to that of PG mRNA) which is inserted into a vector used to transform plants which thereafter have altered ripening properties. The inverted sequence and the PG mRNA form a double-stranded structure which inhibits ex- pression of the PG mRNA. See also AAN80488. (Updated on 25-MAR-2003 to correct PA field.) |
| CC                    |  |
| XX                    |  |
| SQ                    | Sequence 457 AA;   |
| Query Match           | 100.0%; Score 2390; DB 1; Length 457;  |
| Best Local Similarity | 100.0%; Pred. No. 1.1e-193;  |

Matches 457; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 MVIQRNSILLIIIFASSISTCRSNVIDNLFFQVYDNLLEBFPHDFQAYLSYLSKNIE 60
Db 1 MVIQRNSILLIIIFASSISTCRSNVIDNLFFQVYDNLLEBFPHDFQAYLSYLSKNIE 60
QY 61 SNNNIDKVDKNGIKVINLVSFGAKGDKTYDNIAFEQANNEACSSRTVQVFPKKNXYL 120
Db 61 SNNNIDKVDKNGIKVINLVSFGAKGDKTYDNIAFEQANNEACSSRTVQVFPKKNXYL 120
QY 121 LKQITSGPCRSISIVKIFGSLASASKISDYDRRLMTAFDSVONLVVGGGTTNGNGOV 180
Db 121 LKQITSGPCRSISIVKIFGSLASASKISDYDRRLMTAFDSVONLVVGGGTTNGNGOV 180
QY 181 WMPSSCKIKKSLPCRDAPFALTFFWNCNKLKYNLKSNAQOIHKFESCTNVVASNLMIN 240
Db 181 WMPSSCKIKKSLPCRDAPFALTFFWNCNKLKYNLKSNAQOIHKFESCTNVVASNLMIN 240
QY 241 ASAKSPNTDGVHVSNTQYIOISDPTIGTGDPCISIVGSGQVQATNITCGPHGISIGSL 300
Db 241 ASAKSPNTDGVHVSNTQYIOISDPTIGTGDPCISIVGSGQVQATNITCGPHGISIGSL 300
QY 301 GSGNSEAVSVNTVNEAKIIGAENGVRITKMOGSGQASNIKFLANEMODVKYPIIIDON 360
Db 301 GSGNSEAVSVNTVNEAKIIGAENGVRITKMOGSGQASNIKFLANEMODVKYPIIIDON 360
QY 361 YCDRVEPCIQOFSAVQVQNVYENIKGTSATKVAIKFDCSTNPPCGIIMENINLVGSG 420
Db 361 YCDRVEPCIQOFSAVQVQNVYENIKGTSATKVAIKFDCSTNPPCGIIMENINLVGSG 420
QY 421 KPSEATCKNVHFNNAEHVTPHCTSLSEISDEALVNY 457
Db 421 KPSEATCKNVHFNNAEHVTPHCTSLSEISDEALVNY 457

```

RESULT 2

AAP94619  
ID AAP94619 standard; protein; 457 AA.

XX AAP94619;

AC 25-MAR-2003 (revised)

DT 21-JUN-1990 (first entry)

DE Polygalacturonase (PG) cDNA gene product.

XX Polygalacturonase gene; tomato; ds.

OS Lycopersicon esculentum.

PN US4801540-A.

PD 31-JAN-1989.

XX 02-JAN-1987; 87US-00000201.

XX 28-MAR-1986; 86US-00845676.

PR 17-OCT-1986; 86US-00920574.

XX (CALJ ) CALGENE INC.

PA Hiatt WR, Sheehy RE, Shewmaker CK, Kridl JC, Knaut V;

XX WPI; 1989-053640/07.

XX DR N-PSDB; AAN91112.

PT Tomato polygalacturonase gene - used for modulating expression in plant

XX cells or directing expression of heterologous peptide(s).

PS Disclosure; Page; 8pp; English.

CC cDNA of gene may act as a probe to the genomic sequence, provides a means

XX of modulating the production of PG and acts as a source of the transit

XX

XX

XX

CC peptide which may be joined to heterologous peptides directing them to  
cell wall. (Updated on 25-MAR-2003 to correct PA field.)  
XX  
SQ Sequence 457 AA;

Query Match 100.0%; Score 2390; DB 1; Length 457;  
Best Local Similarity 100.0%; Pred. No. 1.1e-193;  
Matches 457; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 MVIQRNSILLIIIFASSISTCRSNVIDNLFFQVYDNLLEBFPHDFQAYLSYLSKNIE 60
Db 1 MVIQRNSILLIIIFASSISTCRSNVIDNLFFQVYDNLLEBFPHDFQAYLSYLSKNIE 60
QY 61 SNNNIDKVDKNGIKVINLVSFGAKGDKTYDNIAFEQANNEACSSRTVQVFPKKNXYL 120
Db 61 SNNNIDKVDKNGIKVINLVSFGAKGDKTYDNIAFEQANNEACSSRTVQVFPKKNXYL 120
QY 121 LKQITSGPCRSISIVKIFGSLASASKISDYDRRLMTAFDSVONLVVGGGTTNGNGOV 180
Db 121 LKQITSGPCRSISIVKIFGSLASASKISDYDRRLMTAFDSVONLVVGGGTTNGNGOV 180
QY 181 WMPSSCKIKKSLPCRDAPFALTFFWNCNKLKYNLKSNAQOIHKFESCTNVVASNLMIN 240
Db 181 WMPSSCKIKKSLPCRDAPFALTFFWNCNKLKYNLKSNAQOIHKFESCTNVVASNLMIN 240
QY 241 ASAKSPNTDGVHVSNTQYIOISDPTIGTGDPCISIVGSGQVQATNITCGPHGISIGSL 300
Db 241 ASAKSPNTDGVHVSNTQYIOISDPTIGTGDPCISIVGSGQVQATNITCGPHGISIGSL 300
QY 301 GSGNSEAVSVNTVNEAKIIGAENGVRITKMOGSGQASNIKFLANEMODVKYPIIIDON 360
Db 301 GSGNSEAVSVNTVNEAKIIGAENGVRITKMOGSGQASNIKFLANEMODVKYPIIIDON 360
QY 361 YCDRVEPCIQOFSAVQVQNVYENIKGTSATKVAIKFDCSTNPPCGIIMENINLVGSG 420
Db 361 YCDRVEPCIQOFSAVQVQNVYENIKGTSATKVAIKFDCSTNPPCGIIMENINLVGSG 420
QY 421 KPSEATCKNVHFNNAEHVTPHCTSLSEISDEALVNY 457
Db 421 KPSEATCKNVHFNNAEHVTPHCTSLSEISDEALVNY 457

```

RESULT 3

AAR32107  
ID AAR32107 standard; protein; 457 AA.

XX AAR32107;

AC 27-AUG-2003 (revised)

DT 25-MAR-2003 (revised)

DT 05-JUL-1993 (first entry)

DE Polygalacturonase.

XX Polygalacturonase; PG; pectin esterase; PG; expression regulation;

XX fruit softening enzymes; flowering plants; fruiting plants;

XX antiense RNA.

OS Lycopersicon esculentum.

XX EP532060-A1.

XX 17-MAR-1993.

XX 06-NOV-1987; 92EP-00117411.

XX 11-NOV-1986; 86GB-00026879.

XX (ICIL ) IMPERIAL CHEM IND PLC.

XX (ZENNE ) ZENNECA LTD.

XX Bridges IG, Grierson D, Schuch WM;

XX

DR WPI; 1993-087084/11.  
 DR N-PSDB; AAC8415.  
 PT Recombinant DNA for flowering and fruiting plants e.g. tomatoes ripening  
 PT control - comprises base sequence for transcription contg. inverted  
 PT sequence of bases complementary to bases in anti sense ribonucleic acid  
 PT encoding softening enzymes, or gene expression regulation.  
 XX  
 PS Example 12; Fig 1; 20pp; English.  
 CC This is the sequence of polygalacturonase from clone pTOM6. The clone was  
 CC used to isolate the PG promoter in the construction of a vector encoding  
 CC antisense RNA to the PG cDNA and PG gene. This would be useful to  
 CC regulate the expression of the fruit softening enzymes in flowering and  
 CC fruiting plants. Such antisense RNA would delay fruit softening. (Updated  
 CC on 25-MAR-2003 to correct PN field.) (Updated on 25-MAR-2003 to correct  
 CC PE field.) (Updated on 25-MAR-2003 to correct PA field.) (Updated on 27-  
 CC AUG-2003 to correct OS field.)  
 CC  
 XX  
 SQ Sequence 457 AA;

Query Match 100.0%; Score 2390; DB 2; Length 457;  
 Best Local Similarity 100.0%; Pred. No. 1.1e-193;  
 Matches 457; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MVIORNSILLIIIFASSISTCRSNVIDNLFKQVNDNIEQEPFADFOAYLSLSKNIE 60  
 DB 1 MVIORNSILLIIIFASSISTCRSNVIDNLFKQVNDNIEQEPFADFOAYLSLSKNIE 60  
 QY 61 SNNNIDKVDKNGIKVINVLSPFGAKGDKTYDNIAPFOAMWACSSRTPVQFVVPKKNYL 120  
 DB 61 SNNNIDKVDKNGIKVINVLSPFGAKGDKTYDNIAPFOAMWACSSRTPVQFVVPKKNYL 120  
 QY 121 LKQITFSGPCRSSISVKIFGSLFASASKISDYKDRRLMIAFDSVQNLVVGSGGTINGNGQV 180  
 DB 121 LKQITFSGPCRSSISVKIFGSLFASASKISDYKDRRLMIAFDSVQNLVVGSGGTINGNGQV 180  
 QY 181 WMPSSCKINKSLPCRDAPFALTFWNCNKLKVNLLKSKNAQOIHFKPESCTNVVASNLMIN 240  
 DB 181 WMPSSCKINKSLPCRDAPFALTFWNCNKLKVNLLKSKNAQOIHFKPESCTNVVASNLMIN 240  
 QY 241 ASAKSPNTDGVHNSNTQYIQISDTIIGTGDDCISIVSGSQNVQATNITCGPGHGISISGL 300  
 DB 241 ASAKSPNTDGVHNSNTQYIQISDTIIGTGDDCISIVSGSQNVQATNITCGPGHGISISGL 300  
 QY 301 GSGNSEAYVSNVTYVNEAKIIGAENGVRITKWQSGGQASNIKFLANVEMQDVKYPITIDON 360  
 DB 301 GSGNSEAYVSNVTYVNEAKIIGAENGVRITKWQSGGQASNIKFLANVEMQDVKYPITIDON 360  
 QY 361 YCDRVEPCIQOFSAVQKVVYENIKGTSATKVAIKFDCSTNPFCEGIIMENINLVGSESG 420  
 DB 361 YCDRVEPCIQOFSAVQKVVYENIKGTSATKVAIKFDCSTNPFCEGIIMENINLVGSESG 420  
 QY 421 KPSEATCKNVHFNNAEHVTPHCTSLSEISDEALLYNY 457  
 DB 421 KPSEATCKNVHFNNAEHVTPHCTSLSEISDEALLYNY 457

RESULT 4  
 AAB48338  
 ID AAB48338 standard; protein; 457 AA.  
 XX  
 AC AAB48338;  
 XX  
 DT 20-APR-2001 (first entry)  
 DE Tomato polygalacturonase (PG) enzyme.  
 XX  
 KM Pectin; pectin methylase; PME; polygalacturonase; PG; tomato; pTOM6;  
 KM food product; yogurt; milk; fruit juice; whey drink; de-esterification.  
 XX  
 OS Lycopersicon esculentum.  
 XX

PN WO200078982-A1.  
 PD 28-DEC-2000.  
 XX  
 XX 15-JUN-2000; 2000WO-1B000869.  
 XX  
 XX 17-JUN-1999; 99GB-00014209.  
 XX  
 PA (DANI-) DANISCO AS.  
 XX  
 PI Christensen TME, Kreiberg JD;  
 XX  
 DR WPI; 2001-091573/10.  
 DR N-PSDB; AAC84653.  
 XX  
 PT Modifying pectin, for foodstuffs preparation, involves transforming host  
 PT having pectin methylase (PME) and polygalacturonase (PG) activity by  
 PT silencing PG activity, to increase PME to PG ratio.  
 XX  
 PS Disclosure; Fig 1; 78pp; English.

Query Match 100.0%; Score 2390; DB 4; Length 457;  
 Best Local Similarity 100.0%; Pred. No. 1.1e-193;  
 Matches 457; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MVIORNSILLIIIFASSISTCRSNVIDNLFKQVNDNIEQEPFADFOAYLSLSKNIE 60  
 DB 1 MVIORNSILLIIIFASSISTCRSNVIDNLFKQVNDNIEQEPFADFOAYLSLSKNIE 60  
 QY 61 SNNNIDKVDKNGIKVINVLSPFGAKGDKTYDNIAPFOAMWACSSRTPVQFVVPKKNYL 120  
 DB 61 SNNNIDKVDKNGIKVINVLSPFGAKGDKTYDNIAPFOAMWACSSRTPVQFVVPKKNYL 120  
 QY 121 LKQITFSGPCRSSISVKIFGSLFASASKISDYKDRRLMIAFDSVQNLVVGSGGTINGNGQV 180  
 DB 121 LKQITFSGPCRSSISVKIFGSLFASASKISDYKDRRLMIAFDSVQNLVVGSGGTINGNGQV 180  
 QY 181 WMPSSCKINKSLPCRDAPFALTFWNCNKLKVNLLKSKNAQOIHFKPESCTNVVASNLMIN 240  
 DB 181 WMPSSCKINKSLPCRDAPFALTFWNCNKLKVNLLKSKNAQOIHFKPESCTNVVASNLMIN 240  
 QY 241 ASAKSPNTDGVHNSNTQYIQISDTIIGTGDDCISIVSGSQNVQATNITCGPGHGISISGL 300  
 DB 241 ASAKSPNTDGVHNSNTQYIQISDTIIGTGDDCISIVSGSQNVQATNITCGPGHGISISGL 300  
 QY 301 GSGNSEAYVSNVTYVNEAKIIGAENGVRITKWQSGGQASNIKFLANVEMQDVKYPITIDON 360  
 DB 301 GSGNSEAYVSNVTYVNEAKIIGAENGVRITKWQSGGQASNIKFLANVEMQDVKYPITIDON 360  
 QY 361 YCDRVEPCIQOFSAVQKVVYENIKGTSATKVAIKFDCSTNPFCEGIIMENINLVGSESG 420  
 DB 361 YCDRVEPCIQOFSAVQKVVYENIKGTSATKVAIKFDCSTNPFCEGIIMENINLVGSESG 420  
 QY 421 KPSEATCKNVHFNNAEHVTPHCTSLSEISDEALLYNY 457  
 DB 421 KPSEATCKNVHFNNAEHVTPHCTSLSEISDEALLYNY 457

## RESULT 5

ABR43936  
ID ABR43936 standard; protein; 456 AA.  
XX  
AC ABR43936;  
XX  
DT 11-AUG-2003 (first entry)  
XX  
DE Tomato endopolygalacturonase enzyme.  
XX  
DE Tomato endopolygalacturonase enzyme.  
XX  
KW Plant; pectin transferase; xyloglucan; polylysine; heparin;  
XX anticoagulant; endopolygalacturonase; enzyme; tomato.  
XX Lycopersicon esculentum.  
OS  
PN WO2003017950-A2.  
XX  
PD 06-MAR-2003.  
XX  
PF 03-SEP-2002; 2002WO-US028066.  
XX  
PR 31-AUG-2001; 2001US-0316777P.  
XX  
PA (UYGE-) UNIV GEORGIA RES FOUND INC.  
XX  
PI Albersheim P, Djelineo-Albersheim I, Davill A;  
XX WPI; 2003-421044/39.  
XX  
PT Formation of ester or amide bond between ester and alcohol or amine  
XX involves treating the ester or its acid or salt with plant pectin  
XX transferase synthase in the presence of alcohol or amine.  
PS Disclosure; Fig 18; 67pp; English.  
XX  
CC The invention relates to forming an ester or amide bond between monomeric  
CC or polymeric ester or its acid or salt and monomeric or polymeric alcohol  
CC or amine by treating the ester or its acid or salt with a plant pectin  
CC transferase synthase in the presence of alcohol or amine under conditions  
CC to form ester or amide bond. The method is useful for forming an ester or  
CC amide bond between monomeric or polymeric ester or its acid or salt (e.g.  
CC homogalacturonan) and monomeric or polymeric alcohol or amine. It is  
CC useful for producing pectin-based polymers e.g., xyloglucan or D- or L-  
CC polylysine useful for the slow release of compounds in the body e.g.,  
CC heparin with anticoagulant or other pharmaceutical properties. The  
CC present sequence represents an endopolygalacturonase enzyme from tomato  
XX  
SQ Sequence 456 AA;

Query Match 97.8%; Score 2338.5; DB 6; Length 456;  
Best Local Similarity 98.9%; Pred. No. 2.6e-189;  
Matches 452; Conservative 0; Mismatches 4; Indels 1; Gaps 1;  
QY 1 MWQRNSLILILIFASSISTCRSNVDNLFKQVVDNLLEQFAHDFQAYLSYLSKNIE 60  
DB 1 MWQRNSLILILIFASSISTCRSNVDNLFKQVVDNLLEQFAHDFQAYLSYLSKNIE 60  
QY 61 SNNNIDKVDKNGIKIVIVLSFGAKGDKTYDIAFEQAWNEACSSRTPVQVVKNNYVL 120  
DB 61 SNNNIDKVDKNGIKIVIVLSFGAKGDKTYDIAFEQAWNEACSSRTPVQVVKNNYVL 120  
QY 121 LKQITSGPCRSISYKIFGSLASASKISDYKDRRLMIAFDSQVNLVVGSGGTINGNGOV 180  
DB 121 LKQITSGPCRSISYKIFGSLASASKISDYKDRRLMIAFDSQVNLVVGSGGTINGNGOV 180  
QY 181 WPPSSCKIKNSLPCRDAPALTFFMCKNLKNNKSKNAQOIHIFKESCTNVVASLMLIN 240  
DB 181 WPPSSCKIKNSLPCRDAPALTFFMCKNLKNNKSKNAQOIHIFKESCTNVVASLMLIN 240  
QY 241 ASAKSPNTDGVAVSNTQYITQISDTTITGTPDCISIVSGSONVOATMTTCGPHGISISGL 300  
DB 241 ASAKSPNTDGVAVSNTQYITQISDTTITGTPDCISIVSGSONVOATMTTCGPHGISISGL 300

QY 301 GSGNSEAYVSNVTNVEAKIIGAENGVRIKTQGGSGQASNIKPLANVEMQDVKPIIIDON 360  
DB 301 GSGNSEAYVSNVTNVEAKIIGAENGVRIKTQGGSGQASNIKPLANVEMQDVKPIIIDON 360  
QY 361 YCDRVEPCIQPSAIVQVKNVYVENIKGTSATKVAIKFDCSTNPPCEGIIMENINLVGSG 420  
DB 361 YCDRVEPCIQPSAIVQVKNVYVENIKGTSATKVAIKFDCSTNPPCEGIIMENINLVGSG 420  
QY 421 KPESEATCGVFNNAEHVTPHCTSLSEISFDEALLNY 457  
DB 421 KPESEATCGVFNNAEHVTPHCTSLSEISFDEALL-NY 456

## RESULT 6

ABB92243  
ID ABB92243 standard; protein; 438 AA.  
XX  
AC ABB92243;  
XX  
DT 31-MAY-2002 (first entry)  
XX  
DE Herbicidally active polypeptide SEQ ID NO 1454.  
XX  
KW Herbicidal; plant; agriculture; herbicide.  
XX  
OS Arabidopsis thaliana.  
XX  
PN WO200210210-A2.  
XX  
PD 07-FEB-2002.  
XX  
PF 28-AUG-2001; 2001WO-EP009892.  
XX  
PR 28-AUG-2001; 2001WO-EP009892.  
XX  
PA (FARB) BAYER AG.  
XX  
PI Tiejien K, Weidner M;  
XX WPI; 2002-269010/31.  
XX  
PT Identifying plant target proteins for herbicidally active compounds,  
XX comprising aligning and comparing nucleic acid or amino acid sequences  
XX from plant with nucleic acid or amino acid sequences from non-plant  
XX organisms.  
PS Claim 5; SEQ ID NO 1454; 261pp + Sequence Listing; English.  
XX  
CC The invention relates to identifying target proteins (ABB90790-ABB94016)  
CC for herbicidally active compounds, comprising aligning and comparing  
CC nucleic acid or amino acid sequences from plant with nucleic acid or  
CC amino acid sequences from non-plant organisms using suitable search  
CC parameters, where plant sequences having an E-value greater by a factor  
CC of 3 than the E-value of most similar non-plant sequences are selected.  
CC The polypeptides or nucleic acids encoding them are useful for  
CC herbicides  
XX  
SQ Sequence 438 AA;

Query Match 49.6%; Score 1185; DB 5; Length 438;  
Best Local Similarity 51.9%; Pred. No. 1.6e-91;  
Matches 230; Conservative 76; Mismatches 121; Indels 16; Gaps 5;  
QY 8 ILLILIFASSISTCRSNV---IDNLFKO---VYDNLLEQFAHDFQAYLSYLSKNIES 61  
DB 5 ILLILIFASSISTCRSNV---IDNLFKO---VYDNLLEQFAHDFQAYLSYLSKNIES 61  
QY 62 NNNIDKVDKNGIKIVIVLSFGAKGDKTYDIAFEQAWNEACSSRTPVQVVKNNYVL 121  
DB 57 KVRHGYAPRRSPRFNNVTGGAAGND-DSKAFMZWAAACSSGTGIYIVAPKRDVMTL 115  
QY 122 KQITSGPCRSISYKIFGSLASASKISDYKDRRLMIAFDSQVNLVVGSGGTINGNGOV 181

Db 116 KATTFGCPCKSLIFLYGRIEAMENPSDYKERHWHVFENNVNLTREVGGRIDGCHIT 175  
Qy 182 WPSCKINSLPBRDAPALTPMCKNKLKNNLSKNAQOIHKFESCTNVASNMINA 241  
Db 176 WPKSCINPOLPLIGATVATFVCCNNLRVSNRLNNAQOMHLTFDDCKRVKLNLMVTS 235  
Qy 242 SAKSPNTDGVHVSANTQYIOISPTIIGTGDPCISIVSGSQNVQATNTTCGPHGISISGL 301  
Db 236 PADSPITDGIHVGSTONILIODSVRTGDDCISIVSGSENVRAITGTCGPHGISISGL 295  
Qy 302 SGNSSEAVSNVNTVEAKIIGANGVRIKTIWQSGGASNIKFLNVEMQDVYPIIIDONY 361  
Db 296 EDNSEAVSNVNVNKAITLIGTTNGVRIKTIWQGGHMAKNIIFODIMKNTNPIIINODY 355  
Qy 362 CDRVEPCIQFSAVQVKNVYENIKGSATKVAIKFDCSNFPCGEGIMENITLVGESGK 421  
Db 356 CDRVEACPCKSAVQVSNVLYKNIQGTSSRPVIAKFCVSKNIPCRGISMQNVKLVDTQ 415  
Qy 422 P-SEATCKNVHFNNAEHVTPHCT 443  
Db 416 DVSKASCNVKLDTRGNVSPFCT 438

RESULT 7  
AAW98178  
ID AAW98178 standard; protein; 431 AA.  
XX AAW98178;  
XX  
XX  
DT 05-JUL-1999 (first entry)  
DE Anther-specific Esj2a protein.  
XX  
XX  
KW Esj2a gene; promoter; pollen; anther dehiscence; male sterile;  
KM transgenic plant.  
XX  
XX  
OS Arabidopsis thaliana.  
XX  
XX  
PN W09913089-A1.  
PD 18-MAR-1999.  
XX  
XX  
PF 11-SEP-1998; 98WC-GB002752.  
XX  
XX  
PR 11-SEP-1997; 97GB-00019359.  
XX  
XX  
PA (BIOG-) BIOGEMMA UK LTD.  
XX  
XX  
PI Roberts JA, Paul W, Craze M;  
XX  
XX  
DR WPI; 1999-254279/21.  
DR N-PSDB; AAX25011.  
XX  
XX  
PT Generation of male sterile plants by controlling anther dehiscence.  
XX  
XX  
PS Disclosure; Fig 3; 34pp; English.

This protein is encoded by the Esj2a gene (see AAX25011) of Arabidopsis thaliana. The invention relates to the use of the Esj2a promoter to reduce dehiscence and to create male sterile plants for use in hybrid seed production. The promoter is used to drive expression of a further nucleic acid sequence that results in prevention or reduction of anther dehiscence. For example, expression of the Kase barnase causes cell ablation, while expression of a plant hormone alters the developmental fate of a cell. Plants are produced that have phenotypically normal pollen grains, within phenotypically normal anthers, but in which the anthers do not dehisce and thus do not release the pollen grains. The system allows the female to be multiplied with the artificial male sterility gene in the homozygous state, since the female plant produces viable pollen. The system is suited to crops which have high seed multiplication, large amounts of pollen, and/or separate male and female inflorescences. These factors allow for easy collection of pollen from

CC non-dehiscent anthers, facile self-pollination and the minimisation of  
CC the area of plants that have to be self-pollinated manually. Such an  
CC ideal crop is the monocot maize, but the system is also applicable to  
CC e.g. wheat, barley, rice, fodder grass, banana, palm, orchid, tulip,  
CC lily, melon, cucumber, tomato, pepper and willow. It is also useful in  
CC the avoidance or reduction of pollen allergens and may be effective in  
CC the control of asthma caused by pollen release  
XX  
XX  
SQ Sequence 431 AA;  
Query Match 47.2%; Score 1127; DB 2; Length 431;  
Best Local Similarity 49.6%; Pred. No. 1,3e-86;  
Matches 223; Conservative 71; Mismatches 116; Indels 40; Gaps 6;  
Qy 7 SILLIIIFASSISTCRSNVID-----DNLFKQVONILIEQFAHFOAYSLY 55  
Db 9 AVFLCVLMLSLCKALSNVNDGYGHEDGSFESDSLKLKNDVLT-----SLSSD 59  
Qy 56 SKNIESNNNIDKVDKNGIVINVLSPGAKGDKTYDNIAFEQAMNACSSRTPVQFVPR 115  
Db 60 ETLLEAS-----TVGSNFGAKGDKGTDGTFQAFKAMKAKCSTNGVTFLEVR 107  
Qy 116 NKNYLLKQITFGCPCKSSISVKIFGSLNASKISDYKDRMLVAFPSVQNLVVGCG--GT 173  
Db 108 GKTYYLTKSTRFPCKSLRNFPQILGTLASATYKSDYKDNHMLIEDVNNLSIDGSTGI 167  
Qy 174 INGNQVWMPSSCKIKSLPCRDAPALTPMCKNKLKNNLSKNAQOIHKFESCTNV 233  
Db 168 INGNKTMQNSCKIDSKFCTKAPALTLVNLKLNLVKNAQOIHKFESCTNRYE 227  
Qy 234 ASNLMINASAKSPNTDGVHVSNTQYIOISPTIIGTGDPCISIVSGSQNVQATNTTCGPH 293  
Db 228 VSNVEITAPGDSGNTDGHITMTQNRVNSDGTGDDCISIBDGTQNIIFDLTCGPH 287  
Qy 294 GISIGSLGSGNSEAVYNSVNTVEAKIIGANGVRIKTIWQSGGASNIKFLNVEMQDVYK 353  
Db 288 GISIGSLGSDNSKAYVSGINVDGAKFESBNGVRIKTIWQSGGTAKNIFQNRMBNVKN 347  
Qy 354 PIITDQYCDRVEPCIQFSAVQVKNVYENIKGSATKVAIKFDCSNFPCGEGIMENI 413  
Db 348 PIITDQYCDK-DKCDQSAVQVKNVYKNSISATDVAILNCSSEKIPCGIIVLENV 406  
Qy 414 NLVSGSKPSEATCKNVHFNNAEHVTPHCT 443  
Db 407 KIKG-----GTASCKNVXNCGTSPKCS 431

RESULT 8  
AAG41324  
ID AAG41324 standard; protein; 431 AA.  
XX AAG41324;  
XX  
XX  
DT 18-OCT-2000 (first entry)  
DE Arabidopsis thaliana protein fragment SEQ ID NO: 51400.  
XX  
XX  
KW Protein identification; signal transduction pathway; metabolic pathway;  
KM hybridisation assay; genetic mapping; gene expression control; promoter;  
XX  
XX  
OS Arabidopsis thaliana.  
XX  
XX  
PN EP1033405-A2.  
PD 06-SEP-2000.  
XX  
XX  
PF 25-FEB-2000; 2000EP-00301439.  
XX  
XX  
PR 25-FEB-1999; 99US-0121825P.  
PR 05-MAR-1999; 99US-0123180P.  
PR 09-MAR-1999; 99US-0123548P.  
PR 23-MAR-1999; 99US-0125788P.



PR 25-MAR-1999; 99US-0126264P.  
PR 29-MAR-1999; 99US-0126785P.  
PR 01-APR-1999; 99US-0127462P.  
PR 06-APR-1999; 99US-0128234P.  
PR 08-APR-1999; 99US-0128714P.  
PR 16-APR-1999; 99US-0129645P.  
PR 19-APR-1999; 99US-0130077P.  
PR 21-APR-1999; 99US-0130449P.  
PR 23-APR-1999; 99US-0130510P.  
PR 28-APR-1999; 99US-0130891P.  
PR 30-APR-1999; 99US-0131449P.  
PR 04-MAY-1999; 99US-0132048P.  
PR 05-MAY-1999; 99US-0132480P.  
PR 06-MAY-1999; 99US-0132485P.  
PR 06-MAY-1999; 99US-0132486P.  
PR 07-MAY-1999; 99US-0132487P.  
PR 11-MAY-1999; 99US-0134256P.  
PR 14-MAY-1999; 99US-0134263P.  
PR 14-MAY-1999; 99US-0134218P.  
PR 14-MAY-1999; 99US-0134219P.  
PR 14-MAY-1999; 99US-0134221P.  
PR 18-MAY-1999; 99US-0134370P.  
PR 19-MAY-1999; 99US-0134768P.  
PR 20-MAY-1999; 99US-0135124P.  
PR 21-MAY-1999; 99US-0135124P.  
PR 24-MAY-1999; 99US-0135353P.  
PR 25-MAY-1999; 99US-0135629P.  
PR 27-MAY-1999; 99US-0136021P.  
PR 28-MAY-1999; 99US-0136392P.  
PR 01-JUN-1999; 99US-0136782P.  
PR 03-JUN-1999; 99US-0137222P.  
PR 04-JUN-1999; 99US-0137528P.  
PR 07-JUN-1999; 99US-0137502P.  
PR 08-JUN-1999; 99US-0137724P.  
PR 10-JUN-1999; 99US-0138094P.  
PR 14-JUN-1999; 99US-0138540P.  
PR 16-JUN-1999; 99US-0138647P.  
PR 17-JUN-1999; 99US-0139119P.  
PR 18-JUN-1999; 99US-0139452P.  
PR 18-JUN-1999; 99US-0139453P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 21-JUN-1999; 99US-0139750P.  
PR 21-JUN-1999; 99US-0139763P.  
PR 22-JUN-1999; 99US-0139817P.  
PR 23-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 24-JUN-1999; 99US-0140659P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140823P.  
PR 30-JUN-1999; 99US-0140916P.  
PR 01-JUL-1999; 99US-0141287P.  
PR 02-JUL-1999; 99US-0141842P.  
PR 06-JUL-1999; 99US-0142154P.  
PR 08-JUL-1999; 99US-0142055P.  
PR 08-JUL-1999; 99US-0142390P.  
PR 09-JUL-1999; 99US-0142803P.  
PR 12-JUL-1999; 99US-0142920P.  
PR 13-JUL-1999; 99US-0142977P.  
PR 14-JUL-1999; 99US-0143542P.  
PR 15-JUL-1999; 99US-0143624P.  
PR 16-JUL-1999; 99US-0144005P.  
PR 99US-0144085P.  
  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 20-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 21-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 22-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 23-JUL-1999; 99US-0145089P.  
PR 23-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 26-JUL-1999; 99US-014524P.  
PR 27-JUL-1999; 99US-014524P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 28-JUL-1999; 99US-0145918P.  
PR 02-AUG-1999; 99US-0145919P.  
PR 02-AUG-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 05-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 06-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 09-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147493P.  
PR 10-AUG-1999; 99US-0147935P.  
PR 11-AUG-1999; 99US-0148317P.  
PR 12-AUG-1999; 99US-0148317P.  
PR 13-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 16-AUG-1999; 99US-0148684P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149368P.  
PR 20-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 23-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 25-AUG-1999; 99US-0149930P.  
PR 26-AUG-1999; 99US-0150566P.  
PR 27-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 30-AUG-1999; 99US-0151080P.  
PR 31-AUG-1999; 99US-0151303P.  
PR 01-SEP-1999; 99US-0151438P.  
PR 07-SEP-1999; 99US-0153070P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154018P.  
PR 20-SEP-1999; 99US-0154039P.  
PR 22-SEP-1999; 99US-0154779P.  
PR 23-SEP-1999; 99US-0155139P.  
PR 24-SEP-1999; 99US-0155486P.  
PR 28-SEP-1999; 99US-0155659P.  
PR 29-SEP-1999; 99US-0156596P.  
PR 04-OCT-1999; 99US-0157117P.  
PR 05-OCT-1999; 99US-0157153P.  
PR 06-OCT-1999; 99US-0157865P.  
PR 07-OCT-1999; 99US-0158029P.

```

PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 13-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161922P.
PR 28-OCT-1999; 99US-0161933P.
PR 29-OCT-1999; 99US-0162142P.

```

Query Match 47.2%; Score 1127; DB 3; Length 431;

Best Local Similarity 49.6%; Pred. No. 1,3e-86; Matches 223; Conservative 71; Mismatches 116; Indels 40; Gaps 6;

```

QY 7 SILLIIIFASSISTCRSNVID-----DNLFKQVYDNLIOEFADHFOAYLSYL 55
DB 9 AVFLCVLMLSLCKALSSNVDDGYGHEDGSFESDSLKLNDDVL-----SLISSD 59
QY 56 SKNIESNNNDKDKNGKIVINVLSPGAKDGGKTYDNIAFEQAMNEACSRTPVQFVPR 115
DB 60 ETTLEAS-----TVSVSNFGAKDGGKTDYQAFKXKWKACSTNGVTTFLVPR 107
QY 116 NKRYLLKQITFGSPCRSSISVKIFGSLFASSTKIDYKDRMLWIAFDSVONLVVGGG--GT 173
DB 108 GKTYLLKSTRFRGPCSLRNFOILGTLSASTKSDYKDKNHWLILDEVNKLSDIGSGTGI 167
QY 174 INNGGVWPPSSCKINKSLPCRDAPALTFTWNCNKLKVNNLKSKNAQOIHIKIESCTNVV 233
DB 168 INNGKTMWNSCKIDSKRCKTAPALTLYNLKLNVLKVLRYKMAQOIOISIEKCNKVE 227
QY 234 ASNLMTINASAKSPNTDGVHVSNTQYQISPTIIGTGDDCISIVSGSONVATNITGSPGH 293
DB 228 VSNVEITAPDSDSNTDGIHTNTQNRVNSDITGDDCISIEDGTQNLQIPDLTGPCH 287
QY 294 GISIGSLGSGNSBAVYSNVTVNEAKIIGAENGVRITKMOGSGOASNIKELANEMODVYK 353
DB 288 GISIGSLGSDNSKAYVSGINVDGAKFESDNGVRITKYOGSGTAKNIKQNIIRMEVKN 347
QY 354 PIIIDQNYCDRVEPCIQFSAVOVKVNVYENIKGTSATKVAIKFDCSTNPPCGIIMENT 413
DB 348 PIIIDQDYCDK-DKCEQDSBAVOVKVNVYENIKGTSATVAITLNCSEKYPCCGIVLENV 406
QY 414 NLVGESGKPSSEATCKNVHFNNAEHTVPHCT 443
DB 407 KING-----GTASCKNANVKNQGVSPKCS 431

```

RESULT 9  
 ABB92704 standard; protein; 431 AA.  
 XX  
 AC ABB92704;

```

XX 31-MAY-2002 (first entry)
DT Herbicidally active polypeptide SEQ ID NO 1915.
XX
DE Herbicidally active polypeptide SEQ ID NO 1915.
XX
KW Herbicidally active polypeptide; herbicide.
XX
OS Arabidopsis thaliana.
XX
PN WO200210210-A2.
XX
PD 07-FEB-2002.
XX
PF 28-AUG-2001; 2001WO-EP009892.
XX
PR 28-AUG-2001; 2001WO-EP009892.
XX
PA (FARB) BAYER AG.
XX
PI Tietjen K, Weidler M;
XX
PI WPI; 2002-269010/31.
XX
DR Identifying plant target proteins for herbicidally active compounds,
XX comprising aligning and comparing nucleic acid or amino acid sequences
XX from plant with nucleic acid or amino acid sequences from non-plant
XX organisms.
XX
PS Claim 5; SEQ ID NO 1915; 261pp + Sequence Listing; English.
XX
CC The invention relates to identifying target proteins (ABB90790-ABB94016)
XX for herbicidally active compounds, comprising aligning and comparing
XX nucleic acid or amino acid sequences from plant with nucleic acid or
XX amino acid sequences from non-plant organisms using suitable search
XX parameters, where plant sequences having an E-value greater by a factor
XX of 3 than the E-value of most similar non-plant sequences are selected.
XX The polypeptides or nucleic acids encoding them are useful for
XX identifying modulators. The identified modulators are useful as
XX herbicides
XX
SQ Sequence 431 AA;

```

Query Match 47.2%; Score 1127; DB 5; Length 431;  
 Best Local Similarity 49.6%; Pred. No. 1,3e-86;  
 Matches 223; Conservative 71; Mismatches 116; Indels 40; Gaps 6;

```

QY 7 SILLIIIFASSISTCRSNVID-----DNLFKQVYDNLIOEFADHFOAYLSYL 55
DB 9 AVFLCVLMLSLCKALSSNVDDGYGHEDGSFESDSLKLNDDVL-----SLISSD 59
QY 56 SKNIESNNNDKDKNGKIVINVLSPGAKDGGKTYDNIAFEQAMNEACSRTPVQFVPR 115
DB 60 ETTLEAS-----TVSVSNFGAKDGGKTDYQAFKXKWKACSTNGVTTFLVPR 107
QY 116 NKRYLLKQITFGSPCRSSISVKIFGSLFASSTKIDYKDRMLWIAFDSVONLVVGGG--GT 173
DB 108 GKTYLLKSTRFRGPCSLRNFOILGTLSASTKSDYKDKNHWLILDEVNKLSDIGSGTGI 167
QY 174 INNGGVWPPSSCKINKSLPCRDAPALTFTWNCNKLKVNNLKSKNAQOIHIKIESCTNVV 233
DB 168 INNGKTMWNSCKIDSKRCKTAPALTLYNLKLNVLKVLRYKMAQOIOISIEKCNKVE 227
QY 234 ASNLMTINASAKSPNTDGVHVSNTQYQISPTIIGTGDDCISIVSGSONVATNITGSPGH 293
DB 228 VSNVEITAPDSDSNTDGIHTNTQNRVNSDITGDDCISIEDGTQNLQIPDLTGPCH 287
QY 294 GISIGSLGSGNSBAVYSNVTVNEAKIIGAENGVRITKMOGSGOASNIKELANEMODVYK 353
DB 288 GISIGSLGSDNSKAYVSGINVDGAKFESDNGVRITKYOGSGTAKNIKQNIIRMEVKN 347
QY 354 PIIIDQNYCDRVEPCIQFSAVOVKVNVYENIKGTSATKVAIKFDCSTNPPCGIIMENT 413
DB 348 PIIIDQDYCDK-DKCEQDSBAVOVKVNVYENIKGTSATVAITLNCSEKYPCCGIVLENV 406

```

OY 414 NLVSGSKPSEATCKNVFNNAEHVDPCT 443  
Db 407 KIRG-----GPAACKMANVKNQGVSPKCS 431

## RESULT 10

AAAG41323  
ID AAAG41323 standard; protein; 463 AA.

AC AAAG41323;

DT 18-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 51399.

KM Protein identification; signal transduction pathway; metabolic pathway;  
hybridisation assay; genetic mapping; gene expression control; promoter;

OS Arabidopsis thaliana.

PN EPI033405-A2.

PD 06-SEP-2000.

PF 25-FEB-2000; 2000EP-00301439.

XX 25-FEB-1999; 99US-0121825P.

PR 05-MAR-1999; 99US-0123180P.

PR 09-MAR-1999; 99US-0123548P.

PR 23-MAR-1999; 99US-0125788P.

PR 25-MAR-1999; 99US-0125264P.

PR 29-MAR-1999; 99US-0126785P.

PR 01-APR-1999; 99US-0127462P.

PR 06-APR-1999; 99US-0128714P.

PR 16-APR-1999; 99US-0129845P.

PR 19-APR-1999; 99US-0130077P.

PR 21-APR-1999; 99US-0130449P.

PR 23-APR-1999; 99US-0130510P.

PR 28-APR-1999; 99US-0130891P.

PR 30-APR-1999; 99US-0132404P.

PR 04-MAY-1999; 99US-0132484P.

PR 05-MAY-1999; 99US-0132485P.

PR 06-MAY-1999; 99US-0132486P.

PR 07-MAY-1999; 99US-0132487P.

PR 11-MAY-1999; 99US-0132863P.

PR 14-MAY-1999; 99US-0134218P.

PR 14-MAY-1999; 99US-0134219P.

PR 18-MAY-1999; 99US-0134370P.

PR 19-MAY-1999; 99US-0134768P.

PR 20-MAY-1999; 99US-0134941P.

PR 21-MAY-1999; 99US-0135353P.

PR 24-MAY-1999; 99US-0135629P.

PR 25-MAY-1999; 99US-0136021P.

PR 27-MAY-1999; 99US-0136392P.

PR 01-JUN-1999; 99US-0136782P.

PR 03-JUN-1999; 99US-0137222P.

PR 04-JUN-1999; 99US-0137528P.

PR 07-JUN-1999; 99US-0137724P.

PR 08-JUN-1999; 99US-0138094P.

PR 10-JUN-1999; 99US-0138540P.

PR 14-JUN-1999; 99US-0138847P.

PR 16-JUN-1999; 99US-0139119P.

PR 16-JUN-1999; 99US-0139452P.

PR 16-JUN-1999; 99US-0139453P.

PR 17-JUN-1999; 99US-0139492P.

PR 18-JUN-1999; 99US-0139454P.

PR 18-JUN-1999; 99US-0139455P.

PR 18-JUN-1999; 99US-0139456P.

PR 18-JUN-1999; 99US-0139457P.

PR 18-JUN-1999; 99US-0139458P.

PR 18-JUN-1999; 99US-0139459P.

PR 18-JUN-1999; 99US-0139460P.

PR 18-JUN-1999; 99US-0139461P.

PR 18-JUN-1999; 99US-0139462P.

PR 18-JUN-1999; 99US-0139463P.

PR 18-JUN-1999; 99US-0139750P.

PR 21-JUN-1999; 99US-0139763P.

PR 22-JUN-1999; 99US-0139817P.

PR 23-JUN-1999; 99US-0139899P.

PR 23-JUN-1999; 99US-0140353P.

PR 24-JUN-1999; 99US-0140354P.

PR 26-JUN-1999; 99US-0140695P.

PR 29-JUN-1999; 99US-0140823P.

PR 30-JUN-1999; 99US-0140991P.

PR 01-JUL-1999; 99US-0141287P.

PR 02-JUL-1999; 99US-0142154P.

PR 06-JUL-1999; 99US-0142390P.

PR 08-JUL-1999; 99US-0142803P.

PR 09-JUL-1999; 99US-0142920P.

PR 12-JUL-1999; 99US-0143277P.

PR 13-JUL-1999; 99US-0143542P.

PR 14-JUL-1999; 99US-0143624P.

PR 15-JUL-1999; 99US-0144005P.

PR 16-JUL-1999; 99US-0144088P.

PR 19-JUL-1999; 99US-0144086P.

PR 19-JUL-1999; 99US-0144332P.

PR 19-JUL-1999; 99US-0144331P.

PR 19-JUL-1999; 99US-0144333P.

PR 19-JUL-1999; 99US-0144334P.

PR 20-JUL-1999; 99US-0144335P.

PR 20-JUL-1999; 99US-0144352P.

PR 20-JUL-1999; 99US-0144632P.

PR 21-JUL-1999; 99US-0144884P.

PR 21-JUL-1999; 99US-0144814P.

PR 21-JUL-1999; 99US-0145086P.

PR 22-JUL-1999; 99US-0145088P.

PR 22-JUL-1999; 99US-0145085P.

PR 22-JUL-1999; 99US-0145087P.

PR 22-JUL-1999; 99US-0145089P.

PR 23-JUL-1999; 99US-0145192P.

PR 23-JUL-1999; 99US-0145145P.

PR 23-JUL-1999; 99US-0145218P.

PR 26-JUL-1999; 99US-0145274P.

PR 27-JUL-1999; 99US-0145276P.

PR 27-JUL-1999; 99US-0145913P.

PR 27-JUL-1999; 99US-0145918P.

PR 28-JUL-1999; 99US-0145919P.

PR 02-AUG-1999; 99US-0146386P.

PR 02-AUG-1999; 99US-0146388P.

PR 02-AUG-1999; 99US-0146389P.

PR 04-AUG-1999; 99US-0147038P.

PR 04-AUG-1999; 99US-0147204P.

PR 05-AUG-1999; 99US-0147302P.

PR 05-AUG-1999; 99US-0147192P.

PR 06-AUG-1999; 99US-0147260P.

PR 06-AUG-1999; 99US-0147303P.

PR 09-AUG-1999; 99US-0147416P.

PR 09-AUG-1999; 99US-0147493P.

PR 10-AUG-1999; 99US-0147935P.

PR 11-AUG-1999; 99US-0148171P.

PR 12-AUG-1999; 99US-0148319P.

PR 13-AUG-1999; 99US-0148565P.

|    |              |                |
|----|--------------|----------------|
| PR | 13-AUG-1999; | 99US-0148684P. |
| PR | 16-AUG-1999; | 99US-0149366P. |
| PR | 17-AUG-1999; | 99US-0149175P. |
| PR | 18-AUG-1999; | 99US-0149426P. |
| PR | 20-AUG-1999; | 99US-0149722P. |
| PR | 20-AUG-1999; | 99US-0149723P. |
| PR | 20-AUG-1999; | 99US-0149929P. |
| PR | 23-AUG-1999; | 99US-0149902P. |
| PR | 23-AUG-1999; | 99US-0149930P. |
| PR | 25-AUG-1999; | 99US-0150566P. |
| PR | 26-AUG-1999; | 99US-0150884P. |
| PR | 27-AUG-1999; | 99US-0151065P. |
| PR | 27-AUG-1999; | 99US-0151066P. |
| PR | 27-AUG-1999; | 99US-0151080P. |
| PR | 30-AUG-1999; | 99US-0151303P. |
| PR | 31-AUG-1999; | 99US-0151438P. |
| PR | 01-SEP-1999; | 99US-0151930P. |
| PR | 07-SEP-1999; | 99US-0152363P. |
| PR | 10-SEP-1999; | 99US-0153070P. |
| PR | 13-SEP-1999; | 99US-0153756P. |
| PR | 15-SEP-1999; | 99US-0154018P. |
| PR | 16-SEP-1999; | 99US-0154039P. |
| PR | 20-SEP-1999; | 99US-0154779P. |
| PR | 22-SEP-1999; | 99US-0155139P. |
| PR | 23-SEP-1999; | 99US-0155486P. |
| PR | 24-SEP-1999; | 99US-0155659P. |
| PR | 26-SEP-1999; | 99US-0156458P. |
| PR | 29-SEP-1999; | 99US-0156596P. |
| PR | 04-OCT-1999; | 99US-0157117P. |
| PR | 05-OCT-1999; | 99US-0157753P. |
| PR | 06-OCT-1999; | 99US-0157865P. |
| PR | 07-OCT-1999; | 99US-0158029P. |
| PR | 08-OCT-1999; | 99US-0158232P. |
| PR | 12-OCT-1999; | 99US-0158369P. |
| PR | 13-OCT-1999; | 99US-0159293P. |
| PR | 13-OCT-1999; | 99US-0159294P. |
| PR | 13-OCT-1999; | 99US-0159295P. |
| PR | 14-OCT-1999; | 99US-0159329P. |
| PR | 14-OCT-1999; | 99US-0159330P. |
| PR | 14-OCT-1999; | 99US-0159331P. |
| PR | 14-OCT-1999; | 99US-0159637P. |
| PR | 14-OCT-1999; | 99US-0159638P. |
| PR | 18-OCT-1999; | 99US-0159584P. |
| PR | 21-OCT-1999; | 99US-0160741P. |
| PR | 21-OCT-1999; | 99US-0160767P. |
| PR | 21-OCT-1999; | 99US-0160768P. |
| PR | 21-OCT-1999; | 99US-0160770P. |
| PR | 21-OCT-1999; | 99US-0160814P. |
| PR | 21-OCT-1999; | 99US-0160815P. |
| PR | 22-OCT-1999; | 99US-0160980P. |
| PR | 22-OCT-1999; | 99US-0160981P. |
| PR | 22-OCT-1999; | 99US-0160982P. |
| PR | 23-OCT-1999; | 99US-0161404P. |
| PR | 25-OCT-1999; | 99US-0161405P. |
| PR | 25-OCT-1999; | 99US-0161406P. |
| PR | 26-OCT-1999; | 99US-0161359P. |
| PR | 26-OCT-1999; | 99US-0161360P. |
| PR | 26-OCT-1999; | 99US-0161361P. |
| PR | 28-OCT-1999; | 99US-0161920P. |
| PR | 28-OCT-1999; | 99US-0161922P. |
| PR | 29-OCT-1999; | 99US-0161931P. |
| PR | 29-OCT-1999; | 99US-0162142P. |

```

Query Match      47.2% Score 1127; DB 3; Length 463;
                  40.6% Pred. No. 1.5e-86;
Best Local Similarity
Matches 223; Conservative 71; Mismatches 116; Indels 40; Gaps 6;

Qy 7 SLLLIIFPSSISGCRSNVID-----DNLFRQYDNIPIEGEFAHDFQALSYL 55
   :::::|||||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 41 AVFLCVLLMLSLCKALSSNVDDGYGHEDGSPESBSLLKLNNDVVL-----SLISSD 91
   :::::|||||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

Qy 56 SKNIESNNNDIKVDNDGIKIVNLVSPGAKGDGKTYDNIAPQAMNEACSSSTPVOFVVPK 115
   :::::|||||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

```

```

Db      92  ETTLEBAS-----TVSVSNFGAKGDKDTDQAFKAKWKKACSTGVTTFVLVPK 139
QY      116  NKNYLLKOTFFSGPCRRSISVAKIPFSLBASCKISIDYKDRRLMIAPDSYQNLVYCGG--GT 173
Db      140  GKTYLLKSTRFRPQCSKLTNRFQILDTLSTASTRSDSKYKKNMMLILEDNNMISIDGSGTGI 199
QY      174  INNGQVWMPSSGCTINKSLPCRDAPTLATFWNCKRLKNKLKSNRAOQIHKFESCTNVV 233
Db      200  INNGKTMWONSCKIDKSRPCTKAPALTALYNLKNLKNLRYKNAQOIQISLKCKRVE 255
QY      234  ASNLMLINAKSASNTDGVHVSNTQYIOISDRTIGTGDDCISIVSAGSQVQATNTTCGPGH 293
Db      260  VSNVEITAPEDSDNTDGIHTNTQNIKRSNSIDIGDDCISIEBGTQNLQIFDLTCGPGH 319
QY      294  GISIGSLGSGNSEAVYSNTTVNEAKIIGAENGVRIKTVOGSGQASNIKPLNEMODVKY 353
Db      320  GISIGSLGDDNSKAYVSGINDGAKFESDNDVRIKTYQGGSGTAHONIKFQINIMENYKN 379
QY      354  PIITIDONYCDRVEPCTIQPFSAVQKRVVYENIKGTSATKVAIAIKDCSTNPFCEGIIENI 413
Db      380  PIITIDQYCDK-DKCEDQESAVQKRVVYKNISGTSATDVAITLNCSEKYPQCGIYLENV 438
QY      414  NLVGSSEKSEBATCKKRVHNNMAEHTPHCT 443
Db      439  KING-----GTASCKNANVKNQTVSPKCS 463

```

|           |  |
|-----------|--|
| RESULT 11 |  |
| AA641325  |  |
| ID        | AA641325 standard; protein; 415 AA.  |
| XX        |  |
| AC        | AA641325;  |
| XX        |  |
| DT        | 18-OCT-2000 (first entry)  |
| XX        |  |
| DE        | Arabidopsis thaliana protein fragment SEQ ID NO: 51401.  |
| XX        |  |
| KW        | Protein identification; signal transduction pathway; metabolic pathway; hydridasation assay; genetic mapping; gene expression control; promoter; termination sequence. |
| XX        |  |
| OS        | Arabidopsis thaliana.  |
| XX        |  |
| PN        | EP1033405-A2.  |
| XX        |  |
| PD        | 06-SEP-2000.   |
| XX        |  |
| PF        | 25-FEB-2000; 2000EP-00301439.  |
| XX        |  |
| PR        | 25-FEB-1999; 99US-0121825P.  |
| PR        | 05-MAR-1999; 99US-0123180P.  |
| PR        | 09-MAR-1999; 99US-0123548P.  |
| PR        | 23-MAR-1999; 99US-0125788P.  |
| PR        | 25-MAR-1999; 99US-0126264P.  |
| PR        | 29-MAR-1999; 99US-0126785P.  |
| PR        | 01-APR-1999; 99US-0127462P.  |
| PR        | 06-APR-1999; 99US-0128234P.  |
| PR        | 08-APR-1999; 99US-0128714P.  |
| PR        | 16-APR-1999; 99US-0129845P.  |
| PR        | 19-APR-1999; 99US-0130077P.  |
| PR        | 21-APR-1999; 99US-0130449P.  |
| PR        | 23-APR-1999; 99US-0130510P.  |
| PR        | 23-APR-1999; 99US-0130891P.  |
| PR        | 28-APR-1999; 99US-0131449P.  |
| PR        | 30-APR-1999; 99US-0132048P.  |
| PR        | 30-APR-1999; 99US-0132407P.  |
| PR        | 04-MAY-1999; 99US-0132484P.  |
| PR        | 05-MAY-1999; 99US-0132485P.  |
| PR        | 06-MAY-1999; 99US-0132486P.  |
| PR        | 06-MAY-1999; 99US-0132487P.  |
| PR        | 07-MAY-1999; 99US-0132863P.  |
| PR        | 11-MAY-1999; 99US-0134256P.  |
| PR        | 14-MAY-1999; 99US-0134218P.  |

PR 14-MAY-1999; 99US-0134219P.  
PR 14-MAY-1999; 99US-0134221P.  
PR 14-MAY-1999; 99US-0134370P.  
PR 18-MAY-1999; 99US-0134768P.  
PR 19-MAY-1999; 99US-0134941P.  
PR 20-MAY-1999; 99US-0135124P.  
PR 21-MAY-1999; 99US-0135353P.  
PR 24-MAY-1999; 99US-0135629P.  
PR 25-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136392P.  
PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 14-JUN-1999; 99US-0138847P.  
PR 16-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 17-JUN-1999; 99US-0139453P.  
PR 18-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141267P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 09-JUL-1999; 99US-0142803P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144634P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145216P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 03-AUG-1999; 99US-0146389P.  
PR 04-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 05-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 06-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 09-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147433P.  
PR 10-AUG-1999; 99US-0147935P.  
PR 11-AUG-1999; 99US-0148171P.  
PR 12-AUG-1999; 99US-0148319P.  
PR 13-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 16-AUG-1999; 99US-0148684P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149903P.  
PR 25-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 30-AUG-1999; 99US-0151080P.  
PR 31-AUG-1999; 99US-0151303P.  
PR 01-SEP-1999; 99US-0151438P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 20-SEP-1999; 99US-0154779P.  
PR 22-SEP-1999; 99US-0155139P.  
PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0155659P.  
PR 28-SEP-1999; 99US-0156458P.  
PR 29-SEP-1999; 99US-0156596P.  
PR 04-OCT-1999; 99US-0157117P.  
PR 05-OCT-1999; 99US-0157753P.  
PR 06-OCT-1999; 99US-0157865P.  
PR 07-OCT-1999; 99US-0158029P.  
PR 08-OCT-1999; 99US-0158232P.  
PR 12-OCT-1999; 99US-0158369P.  
PR 13-OCT-1999; 99US-0159293P.  
PR 13-OCT-1999; 99US-0159294P.  
PR 13-OCT-1999; 99US-0159329P.  
PR 14-OCT-1999; 99US-0159330P.  
PR 14-OCT-1999; 99US-0159331P.  
PR 14-OCT-1999; 99US-0159337P.  
PR 14-OCT-1999; 99US-0159638P.  
PR 18-OCT-1999; 99US-0159584P.  
PR 21-OCT-1999; 99US-0160741P.  
PR 21-OCT-1999; 99US-0160767P.  
PR 21-OCT-1999; 99US-0160768P.  
PR 21-OCT-1999; 99US-0160770P.  
PR 21-OCT-1999; 99US-0160814P.  
PR 21-OCT-1999; 99US-0160815P.  
PR 22-OCT-1999; 99US-0160980P.  
PR 22-OCT-1999; 99US-0160981P.  
PR 22-OCT-1999; 99US-0160989P.



OS Brassica napus; cv. Rafal.  
 XX W09630529-A1.  
 XX 03-OCT-1996.  
 XX 29-MAR-1996; 96MO-GB000757.  
 XX 31-MAR-1995; 95GB-0006684.  
 XX (NICK-) NICKERSON BIOCHEM LTD.  
 XX Roberts JA, Coupe SA, Jenkins ES;  
 XX WPI; 1996-455374/45.  
 DR N-PSDB; AATJ33994.  
 XX  
 PT Control of seed pod dehiscence - using polygalacturonase or nucleic acid  
 XX sequences derived from polygalacturonase gene.  
 XX  
 PS Example 1; Fig 1; 36pp; English.  
 XX  
 CC A polygalacturonase (PG) (AAW04248) of oilseed rape cv. Rafal seed pods  
 CC is useful for controlling dehiscence. Manipulation of the enzyme's  
 CC coding for the PG was isolated from a cDNA library of the rape seed pod  
 CC dehiscence zone. PG nucleic acids (patric. antisense) can be used to  
 CC regulate dehiscence in crop plants. (Updated on 16-OCT-2003 to  
 CC standardise OS field)  
 XX  
 SQ Sequence 433 AA;

Query Match 46.2%; Score 1103; DB 2; Length 433;  
 Best Local Similarity 49.8%; Pred. No. 1.4e-84;  
 Matches 221; Conservative 70; Mismatches 125; Indels 28; Gaps 7;  
 QY 7 SILLIIIFASSISTGRNVID-----DNLFKQVYDNIIEQFAHDFQAVLYSTKNIYES 61  
 Db 9 AIFLCVLLMLACCOALSSNVDDGYGHEGDSFET--DSLILKLNDDVLTLSKSDRPTTES 66  
 QY 62 NNNIDKVDKNGIKVINLVLSFGAKGDKTYDNIAREQAMNEACSRTPVQFVPRKNRYLL 121  
 Db 67 S-----TVSVNFGAKGDKTDTQAFKAKMKKACSTNGVTTLIRPKGYLL 114  
 QY 122 KOITSGPCRSISVKTIFGSLFASASKISDY-KDRRLMIAPDSYQNTLVVGGG--GTINGNG 178  
 Db 115 KSIRRGFCCKSLRFQILGTLSASTKSDYNDKNHMLLEDVNNLSIDGSGAGIVDNG 174  
 QY 179 QVWMPSSCKINKSLPCBDAPALTALTFMNCCKYLKYNMLKSKAAQOIHIKFECSCTNVVASNLM 238  
 Db 175 KIMMONSCKIDKSPCTKAPTALTLYNLNLANVKNLRVNAQOIQISIEKNCSDVVKVYK 234  
 QY 239 INASAKSPNTDGVAVNSTOYIQTSDTITIGTDCDISIVGSONVOATNITCGPHGISIG 298  
 Db 235 ITAGDSPNTDGIHIVATKIRISNSDITGDCDISIEDGSONVOINDLTGPHGISIG 294  
 QY 299 SLGSGNSEAVSVNTVNEAKITGAENGVRITKMOGSGSQASNIKFLANVEMODVXPPIIID 358  
 Db 295 SLGDDNSKAVSGINVDATLSETDNGVRIKTYGSGGTAKNIKQNTRMNDVKNPIIID 354  
 QY 359 QNYCDREPCIOQFSAVQVKNVYENIKGTSAKYAIFKDCSTNFPCEGIIMENINLVGE 418  
 Db 355 QNYCDR-DKCEQOBSAVQVNNVYRNIGTSAIDVAIMFNCSVYKPCOGIYLVENVIKIG- 412  
 QY 419 SGKPSATKCNVFNNAEHVTPHC 442  
 Db 413 ----GKASCKNVNVKDKGTVSPKC 432

RESULT 14  
 AA42649  
 ID AA42649 standard; protein; 433 AA.

AC AA42649;  
 XX  
 DT 10-JAN-2000 (first entry)  
 XX  
 DE Brassica napus Sac66 protein putative sequence.  
 XX  
 KW Signal transduction protein; dehiscence; male sterile plant; D22 gene;  
 XX shatter resistance; oilseed rape; Sac66 protein.  
 XX  
 OS Brassica napus.  
 XX  
 XX W09949046-A1.  
 XX  
 XX 30-SEP-1999.  
 XX  
 XX 22-MAR-1999; 99MO-GB000905.  
 XX  
 XX 20-MAR-1998; 98GB-0006113.  
 XX  
 PA (BIOG-) BIOGENMA UK LTD.  
 XX  
 PI Wyatt P, Roberts JA, Whiteclaw C;  
 XX  
 XX WPI; 1999-580449/49.  
 DR N-PSDB; AA222980.  
 XX  
 PT A nucleic acid encoding a signal transduction protein involved in plant  
 XX dehiscence, useful for producing shatter resistant male sterile plants.  
 XX  
 PS Example 6; Fig 15; 71pp; English.  
 XX  
 CC The invention provides a nucleic acid encoding a signal transduction  
 CC protein involved in the process of dehiscence. The nucleic acids and  
 CC proteins are useful for regulating or controlling dehiscence of a pod or  
 CC an anther in a plant, useful in the production of male sterile plants.  
 CC The methods, etc. may be used in production of shatter resistance or  
 CC shatter-delayed plants such as oilseed rape (Brassica napus). The present  
 CC sequence represents a B. napus Sac66 protein putative sequence  
 XX  
 SQ Sequence 433 AA;

Query Match 46.2%; Score 1103; DB 2; Length 433;  
 Best Local Similarity 49.8%; Pred. No. 1.4e-84;  
 Matches 221; Conservative 70; Mismatches 125; Indels 28; Gaps 7;  
 QY 7 SILLIIIFASSISTGRNVID-----DNLFKQVYDNIIEQFAHDFQAVLYSTKNIYES 61  
 Db 9 AIFLCVLLMLACCOALSSNVDDGYGHEGDSFET--DSLILKLNDDVLTLSKSDRPTTES 66  
 QY 62 NNNIDKVDKNGIKVINLVLSFGAKGDKTYDNIAREQAMNEACSRTPVQFVPRKNRYLL 121  
 Db 67 S-----TVSVNFGAKGDKTDTQAFKAKMKKACSTNGVTTLIRPKGYLL 114  
 QY 122 KOITSGPCRSISVKTIFGSLFASASKISDY-KDRRLMIAPDSYQNTLVVGGG--GTINGNG 178  
 Db 115 KSIRRGFCCKSLRFQILGTLSASTKSDYNDKNHMLLEDVNNLSIDGSGAGIVDNG 174  
 QY 179 QVWMPSSCKINKSLPCBDAPALTALTFMNCCKYLKYNMLKSKAAQOIHIKFECSCTNVVASNLM 238  
 Db 175 KIMMONSCKIDKSPCTKAPTALTLYNLNLANVKNLRVNAQOIQISIEKNCSDVVKVYK 234  
 QY 239 INASAKSPNTDGVAVNSTOYIQTSDTITIGTDCDISIVGSONVOATNITCGPHGISIG 298  
 Db 235 ITAGDSPNTDGIHIVATKIRISNSDITGDCDISIEDGSONVOINDLTGPHGISIG 294  
 QY 299 SLGSGNSEAVSVNTVNEAKITGAENGVRITKMOGSGSQASNIKFLANVEMODVXPPIIID 358  
 Db 295 SLGDDNSKAVSGINVDATLSETDNGVRIKTYGSGGTAKNIKQNTRMNDVKNPIIID 354  
 QY 359 QNYCDREPCIOQFSAVQVKNVYENIKGTSAKYAIFKDCSTNFPCEGIIMENINLVGE 418  
 Db 355 QNYCDR-DKCEQOBSAVQVNNVYRNIGTSAIDVAIMFNCSVYKPCOGIYLVENVIKIG- 412





This page Blank (uspto)

This page Blank (uspto)

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Comugen Ltd.

# OM protein - protein search, using sw model

Run on: March 24, 2005, 06:59:05 ; Search time 41 Seconds  
(without alignments)  
1072.465 Million cell updates/sec

Title: US-10-691-374-2

Perfect score: 2390

Sequence: 1 MWIQRNSILLILIIIFASSIS.....VTPHCTSLSEIDELALVNV 457

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database: PIR 79:\*

1: p1r1:\*

2: p1r2:\*

3: p1r3:\*

4: p1r4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score  | Query Match | Length | DB | ID     | Description        |
|------------|--------|-------------|--------|----|--------|--------------------|
| 1          | 2390   | 100.0       | 457    | 2  | A25534 | polygalacturonase  |
| 2          | 1142   | 47.8        | 462    | 1  | S31195 | polygalacturonase  |
| 3          | 1128   | 47.2        | 461    | 2  | T08215 | polygalacturonase  |
| 4          | 1127   | 47.2        | 431    | 2  | T46187 | polygalacturonase  |
| 5          | 1115.5 | 46.7        | 426    | 2  | H84846 | probable polygalac |
| 6          | 1108   | 46.4        | 458    | 2  | S71523 | polygalacturonase  |
| 7          | 1091   | 45.6        | 460    | 2  | T17011 | polygalacturonase  |
| 8          | 898.5  | 37.6        | 459    | 2  | D96833 | hypothetical prote |
| 9          | 840.5  | 35.2        | 468    | 2  | H6728  | probable polygalac |
| 10         | 796.5  | 33.3        | 514    | 2  | JC7100 | polygalacturonase  |
| 11         | 787.5  | 32.9        | 434    | 2  | A96609 | probable polygalac |
| 12         | 787.5  | 32.9        | 514    | 2  | S48720 | Cry j II protein - |
| 13         | 787    | 32.9        | 514    | 2  | JC2498 | second major aller |
| 14         | 764    | 32.0        | 507    | 2  | JC7366 | Jun a 2 protein -  |
| 15         | 749.5  | 31.4        | 1161   | 2  | H86368 | protein F28C11.9 [ |
| 16         | 741.5  | 31.0        | 491    | 2  | B86155 | probable polygalac |
| 17         | 736    | 30.8        | 423    | 2  | T08213 | polygalacturonase  |
| 18         | 728    | 30.5        | 492    | 2  | D84871 | probable polygalac |
| 19         | 727    | 30.4        | 392    | 2  | S57806 | polygalacturonase  |
| 20         | 718.5  | 30.1        | 395    | 2  | T05906 | probable polygalac |
| 21         | 718    | 30.0        | 435    | 2  | T48618 | polygalacturonase  |
| 22         | 715    | 29.9        | 392    | 2  | T00669 | probable polygalac |
| 23         | 711.5  | 29.7        | 387    | 2  | T04330 | polygalacturonase  |
| 24         | 710.5  | 29.7        | 365    | 2  | T04319 | polygalacturonase  |
| 25         | 708.5  | 29.6        | 422    | 2  | S34199 | polygalacturonase  |
| 26         | 704.5  | 29.5        | 387    | 2  | T04352 | polygalacturonase  |
| 27         | 704    | 29.5        | 374    | 2  | T04672 | probable polygalac |
| 28         | 704    | 29.5        | 374    | 2  | B85421 | hypothetical prote |
| 29         | 702.5  | 29.4        | 394    | 2  | G86190 |                    |

|    |       |      |     |   |        |                    |
|----|-------|------|-----|---|--------|--------------------|
| 30 | 701.5 | 29.4 | 542 | 2 | H86239 | protein F20B24.8 [ |
| 31 | 698.5 | 29.2 | 452 | 2 | C85024 | probable polygalac |
| 32 | 696   | 29.1 | 384 | 2 | B84871 | probable polygalac |
| 33 | 690   | 28.9 | 540 | 2 | B96631 | probable polygalac |
| 34 | 688.5 | 28.8 | 492 | 2 | C96521 | protein F21D18.18  |
| 35 | 687.5 | 28.8 | 393 | 2 | S40123 | polygalacturonase  |
| 36 | 687   | 28.7 | 383 | 2 | T47809 | polygalacturonase  |
| 37 | 684   | 28.6 | 397 | 2 | F96680 | F5114.10 (imported |
| 38 | 683   | 28.5 | 394 | 2 | F86190 | hypothetical prote |
| 39 | 680.5 | 28.5 | 394 | 2 | T00668 | probable polygalac |
| 40 | 678.5 | 28.4 | 445 | 2 | S34266 | polygalacturonase  |
| 41 | 673.5 | 28.2 | 387 | 2 | T07591 | polygalacturonase  |
| 42 | 673   | 28.2 | 444 | 2 | S34200 | polygalacturonase  |
| 43 | 667   | 27.9 | 407 | 2 | S52006 | polygalacturonase  |
| 44 | 662.5 | 27.7 | 664 | 2 | A84742 | probable polygalac |
| 45 | 659   | 27.6 | 414 | 2 | B85204 | polygalacturonase- |

## ALIGNMENTS

### RESULT 1

A25534 polygalacturonase (EC 3.2.1.15) precursor - tomato  
N:Alternate names: pectinase; poly[1,4-alpha-D-galacturonide]glucanohydrolase  
C:Species: Lycopersicon esculentum (tomato)  
C:Date: 30-Jun-1988 #sequence revision 30-Jun-1988 #text\_change 09-Jul-2004  
C:Accession: A25534; J0156; S06340; S02101; S08504  
R:Giersen, D.; Tucker, G.A.; Keen, J.; Ray, J.; Bird, C.R.; Schuch, W.  
Nucleic Acids Res. 14, 8595-8603, 1986  
A:Title: Sequencing and identification of a cDNA clone for tomato polygalacturonase.  
A:Reference number: A25534; M01D:8706731; PMID:3786135  
A:Accession: A25534  
A:Molecule type: mRNA  
A:Residues: 1-457 <GR1>  
A:Cross-references: UNIPROT:P05117; GB:X04583; NID:g19291; PIDN:CAA28254.1; PID:g19292  
R:DeLlPenna, D.; Bennett, A.B.  
Plant Physiol. 86, 1057-1063, 1988  
A:Title: In vitro synthesis and processing of tomato fruit polygalacturonase.  
A:Reference number: J0156  
A:Accession: J0156  
A:Molecule type: mRNA  
A:Residues: 1-115 <DEL>  
A:Cross-references: GB:M20269; NID:g170470; PIDN:AAA34177.1; PID:g170471  
A:Experimental source: fruit  
R:Sheehy, R.E.; Pearson, J.; Brady, C.J.; Hiatt, W.R.  
Mol. Gen. Genet. 208, 30-36, 1987  
A:Title: Molecular characterization of tomato fruit polygalacturonase.  
A:Reference number: S06340  
A:Accession: S06340  
A:Molecule type: mRNA  
A:Residues: 1-457 <SH>  
A:Cross-references: EMBL:X05656; NID:g19297; PIDN:CAA29148.1; PID:g19298  
A:Note: part of this sequence, including the amino end of the mature protein, was confirm  
R:Rose, R.E.; Houck, C.M.; Monson, E.K.; DeDeus, C.E.; Sheehy, R.E.; Hiatt, W.R.  
Nucleic Acids Res. 16, 7191, 1988  
A:Title: The nucleotide sequence of the 5' flanking region of a tomato polygalacturonase  
A:Reference number: S02101; M01D:88303350; PMID:3405767  
A:Accession: S02101  
A:Molecule type: DNA  
A:Residues: 1-93 <ROS>  
A:Cross-references: EMBL:X07410; NID:g19295; PIDN:CAA30308.1; PID:g19296  
R:Bird, C.R.; Smith, C.J.S.; Ray, J.A.; Moureaux, P.; Bevan, M.W.; Bird, A.S.; Hughes, S.;  
Plant Mol. Biol. 11, 651-662, 1988  
A:Title: The tomato polygalacturonase gene and ripening-specific expression in transgenic  
A:Reference number: S08504  
A:Accession: S08504  
A:Molecule type: DNA  
A:Residues: 1-457 <BR>  
A:Cross-references: EMBL:X14074; NID:g19305; PIDN:CAA32235.1; PID:g295813  
C:Comment: The polygalacturonase, a single cell wall enzyme, is the major enzyme mediactir  
C:Superfamily: polygalacturonase  
C:Keywords: glycosidase; hydrolase; polysaccharide degradation

F:1-25/Domain: signal sequence #status predicted &lt;SIG&gt;

F:26-71/Domain: propeptide #status predicted &lt;PRO&gt;

F:72-444/Product: polygalacturonase 2A #status experimental &lt;MAT&gt;

## Query Match

100.0%; Score 2390; DB 2; Length 457;  
Best Local Similarity 100.0%; Pred. No. 5, 6e-163;  
Matches 457; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVIQNSLILIIIFASSISTCRSNVIDNLFQVYDNLLEQFAHDFQAVLYSLKNIIE 60  
DB 1 MVIQNSLILIIIFASSISTCRSNVIDNLFQVYDNLLEQFAHDFQAVLYSLKNIIE 60  
QY 61 SNNIDVNDKNGIKVINLVSPGAKGDKTYDNIAPQAMNEACSSRTPOGVVPRKNTL 120  
DB 61 SNNIDVNDKNGIKVINLVSPGAKGDKTYDNIAPQAMNEACSSRTPOGVVPRKNTL 120  
QY 121 LKQITFSGCRSSISVKIFGSLBASISKIDYKDRRLWIAPDSYONLVGGGGTINGNGQV 180  
DB 121 LKQITFSGCRSSISVKIFGSLBASISKIDYKDRRLWIAPDSYONLVGGGGTINGNGQV 180  
QY 181 WMPSSCKINKSLPCRDAPATLTFMNCNKLKNNLKSMAQOIHIFESCTNVVASNLMIN 240  
DB 181 WMPSSCKINKSLPCRDAPATLTFMNCNKLKNNLKSMAQOIHIFESCTNVVASNLMIN 240  
QY 241 ASAKSPNTDGVHVSNTQYIQTITITIGTDDCISIVSGSONVOANTICGFGHISIGSL 300  
DB 241 ASAKSPNTDGVHVSNTQYIQTITITIGTDDCISIVSGSONVOANTICGFGHISIGSL 300  
QY 301 GSGNSEAVSNVTVNNEAKITGAENGVRITKMGSGSGQASNIKFVEMQDVKPIIIDON 360  
DB 301 GSGNSEAVSNVTVNNEAKITGAENGVRITKMGSGSGQASNIKFVEMQDVKPIIIDON 360  
QY 361 YCVRVPCITQOFSAVQVKNVYENIKGSAATKVAIKFDCSTNPFCEGIIIMENINLVGSG 420  
DB 361 YCVRVPCITQOFSAVQVKNVYENIKGSAATKVAIKFDCSTNPFCEGIIIMENINLVGSG 420  
QY 421 KSEATCKKVNHNMAEHPTCTLSISEDEALLVNY 457  
DB 421 KSEATCKKVNHNMAEHPTCTLSISEDEALLVNY 457  
QY 421 KSEATCKKVNHNMAEHPTCTLSISEDEALLVNY 457  
DB 421 KSEATCKKVNHNMAEHPTCTLSISEDEALLVNY 457

## RESULT 2

S31195

polygalacturonase (EC 3.2.1.15) - avocado

C:Species: Persea americana (avocado)

C:Date: 10-Sep-1999 #sequence \_revision 10-Sep-1999 #text\_change 09-Jul-2004

R:Accession: S31195; S28072

Plant Mol. Biol. 21, 437-449, 1993

A:Title: Cloning and characterization of avocado fruit mRNAs and their expression during

A:Reference number: S31195; MUID:93184201; PMID:8095163

A:Molecule type: mRNA

A:Residues: 1-462 &lt;DOP&gt;

A:Cross-references: UNIPROT:Q02096; EMBL:X66426

submitted to the EMBL Data Library, May 1992

A:Reference number: S28072

A:Molecule type: mRNA

A:Residues: 1-181, 'S', 183-462 &lt;DOP&gt;

A:Cross-references: EMBL:X66426; NID:G22630; PID:CAA7055.1; PID:G22631

C:Superfamily: polygalacturonase

C:Keywords: glycosidase; hydrolase; polysaccharide degradation

Query Match

Best Local Similarity 47.8%; Score 1142; DB 1; Length 462;  
Matches 224; Conservative 67; Mismatches 106; Indels 8; Gaps 5;QY 28 DDLRFQVYDNLLEQFAHDFQAVLYSLKNIISNNIDVNDKNGIKVINLVSPGAKG 87  
DB 40 DDLRFQVYDNLLEQFAHDFQAVLYSLKNIISNNIDVNDKNGIKVINLVSPGAKG 87

QY 88 KTYNIAPEQAMNEACSSRTPOGVVPRKNTLTKQITFSGCRSSISVKIFGSLBAS 147  
DB 97 -TDTKAFERKAMDACSSSVL--VPEKNTYLLKQITFSGCRSSISVKIFGSLBAS 153  
QY 148 ISDY--KDRRLWIAPDSYONLVGGGGTINGNGQVWMPSSCKINKSLPCRDAPATLTF 205  
DB 154 OSDVGHNKRNITFEEDISNLTLEGGGTINNGSTWDSCKRKSLLPCKSAATLTFRS 213  
QY 206 CKRLKNNLKSMAQOIHIFESCTNVVASNLMINASASPNTDGVHVSNTQYIQTIT 265  
DB 214 CKRLKNNLKSMAQOIHIFESCTNVVASNLMINASASPNTDGVHVSNTQYIQTIT 273  
QY 266 IGTGDDCISIVSGSONVOANTICGFGHISIGSLGSGNSEAVSNVTVNNEAKITGA 325  
DB 274 IGTGDDCISIVSGSONVOANTICGFGHISIGSLGSGNSEAVSNVTVNNEAKITGA 333  
QY 326 VRIKMGSGQASNIKFVEMQDVKPIIIDONVYCDVRVPCITQOFSAVQVKNVYENI 385  
DB 334 LRIKMGSGQASNIKFVEMQDVKPIIIDONVYCDVRVPCITQOFSAVQVKNVYENI 393  
QY 386 KGTSAATVAIKFDCSTNPFCEGIIIMENINLVGSGKSEATCKVNY 430  
DB 394 KGTSAATVAIKFDCSTNPFCEGIIIMENINLVGSGKSEATCKVNY 438

## RESULT 3

T08215

polygalacturonase (EC 3.2.1.15) 3 precursor - muskmelon

C:Species: Cucumis melo (muskmelon)

C:Date: 11-Jun-1999 #sequence \_revision 11-Jun-1999 #text\_change 09-Jul-2004

R:Accession: T08215

Plant Physiol. 117, 363-373, 1998

A:Title: Polygalacturonase gene expression in ripe melon fruit supports a role for polyg

A:Reference number: T16403; MUID:98289082; PMID:9625689

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-461 &lt;HAD&gt;

A:Cross-references: UNIPROT:O81246; EMBL:AF062467; NID:G3320461; PID:G3320462

C:Genetics:

A:Gene: MRC3

C:Superfamily: polygalacturonase

C:Keywords: glycosidase; hydrolase; polysaccharide degradation

F:1-34/Domain: signal sequence #status predicted &lt;SIG&gt;

F:35-461/Product: polygalacturonase #status predicted &lt;MAT&gt;

Query Match

Best Local Similarity 47.2%; Score 1128; DB 2; Length 461;  
Matches 224; Conservative 79; Mismatches 129; Indels 14; Gaps 7;

QY 7 SILIIIIIFASSISTCRSNVIDNLFQVYDNLLEQFAHDFQAVLYSLKNIISNNI 65  
DB 17 TFLFVNVNFDFTSTCFSGYPPVDNPLPSISSGDKEDVGHYHSSSEFSSMLRT--RL 74  
QY 66 DKVDKNGI--KVIVNLSPGAKGDKTYDNIAPQAMNEACSSRTPOGVVPRKNTL 122  
DB 75 EKNVSSPLASPEIRFVNDVYDGMGDEG--DTEAFRETWKDACSTNAT--PLVVCDEYHILK 132  
QY 123 QITFSGCRSSISVKIFGSLBASISKIDY--KDRRLWIAPDSYONLVGGGGTINGNGQV 181  
DB 133 QITFSGCRSSISVKIFGSLBASISKIDY--KDRRLWIAPDSYONLVGGGGTINGNGQV 192  
QY 182 WMPSSCKINKSLPCRDAPATLTFMNCNKLKNNLKSMAQOIHIFESCTNVVASNLMIN 241  
DB 193 WMPSSCKINKSLPCRDAPATLTFMNCNKLKNNLKSMAQOIHIFESCTNVVASNLMIN 252  
QY 242 SAKSPNTDGVHVSNTQYIQTITITIGTDDCISIVSGSONVOANTICGFGHISIGSL 301  
DB 253 SAKSPNTDGVHVSNTQYIQTITITIGTDDCISIVSGSONVOANTICGFGHISIGSL 312  
QY 302 GSGNSEAVSNVTVNNEAKITGAENGVRITKMGSGSGQASNIKFVEMQDVKPIIIDON 361

Db 313 AGKSEAEVSNVVDYTAKEFGSTNGVRIRKTVQGGKGYAQNIIIFONIVMDVNTNPIIINQNY 372  
Qy 362 CDRVEPCIQOFSAVQVKNVYENIKGTSATKVAIKFDCSTNPFCEGIIEMENILV----- 416  
Db 373 CDQKEPCTQADAVANVNMVKNIRGTSASEVAVKFDCSKSVPCQGLIDDLINLVHKGKN 432  
Qy 417 GEGSKPESEATCKNVHFNNAEHVTPHC 442  
Db 433 DDKSAQAEASCKNVKMKNGRVRSPQC 458

## RESULT 4

T46187  
polygalacturonase (EC 3.2.1.15) precursor [similarity] - Arabidopsis thaliana  
N:Alternate names: protein T8H10.110  
C/Species: Arabidopsis thaliana (mouse-ear cress)  
C/Date: 04-Feb-2000 #sequence\_revision 04-Feb-2000 #text\_change 09-Jul-2004  
C/Accession: T46187, T50674  
R/Bones, V.; Rechmann, S.; Borkova, D.; Ansgore, W.; Mewes, H.W.; Lemcke, K.; Mayer, K. H.  
submitted to the Protein Sequence Database, January 2000  
A/Reference number: Z23014  
A/Accession: T46187  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-431 <BEN>  
A/Cross-references: UNIPROT:O23147; EMBL:AL133248; PIDN:CAB66108.1  
A/Experimental source: cultivar Columbia; BAC clone T8H10  
R/Jenkins, E.S.; Roberts, J.A.  
submitted to the EMBL Data Library, December 1997  
A/Description: Dehiscence-related expression of an Arabidopsis thaliana gene encoding a  
A/Reference number: Z25172  
A/Accession: T50674  
A/Status: preliminary; translated from GB/EMBL/DBJ  
A/Molecule type: DNA  
A/Residues: 1-431 <DEN>  
A/Cross-references: EMBL:AF037367; PIDN:AAC98923.1  
A/Experimental source: cultivar Landsberg erecta  
C/Genetic:  
A/Map position: 3  
A/Intons: 85/3; 129/3; 186/3; 193/3; 263/1; 290/2; 326/3; 364/3  
C/Superfamily: polygalacturonase  
C/Keywords: glycosidase; hydrolase  
F/1-27/Domain: signal sequence #status predicted <SIG>  
F/28-431/Product: polygalacturonase #status predicted <MAT>

## Query Match

Best Local Similarity 49.6%; Score 1127; DB 2; Length 431;  
Matches 223; Conservative 71; Mismatches 116; Indels 40; Gaps 6;

Qy 7 SILLIIIFASSISTCRSNVID-----DNLEFKOYVDNILEQFAHDFQAYLSYL 55  
Db 9 AVLCLVLMWLCKLAUSSNVDDGYGHEHDSFESDLSLKNNDVL-----SLISSD 59  
Qy 56 SKNIESNNIDKYDKNGIKYINLVSPGAKGDKTYDNIAEFOAMNEACSSRTPOFVVPK 115  
Db 60 ETTLEAS-----TVSVNFGAKGDKTDQGAKKAKKACSTNGVTFPLVPK 107  
Qy 116 NKNYLLKQITFGSPCRSSISVKIFGSLSEASSKISDYKDRRLMIAFDSVNLVVGCG--GT 173  
Db 108 GKTYLLKSTRFGRPKSLRNFQILGTLASASTKSDYDKKNHMLILEVDNMLSIDGSGTGI 167  
Qy 174 INNGGVWPPSSCKINSLPCRDAPTLTFMNCCKNLKNNLKSNAQOHHIKESCTNVV 233  
Db 168 INNGKTMWNSCKIDSKPCTKAPVALTLYNLKNLVKNLRYKNAQOIQISIECKNVE 227  
Qy 234 ASNLMIINASAKSPNTDGVHVSNTQYIQISDTIIGTGDDCISIVSGSNVQATNITCGPGH 293  
Db 228 VSNVEITRAPDPSRPTDGIHTNTQNIKRVNSDIDGTGDCISIEBGTQNLIDPLDTGPGH 287  
Qy 294 GISIGISGSGNSEAVYNSVTNVEAKITGAENGVRIRKTVQGGSGOASNIKFLVEMODVKY 353  
Db 288 GISIGISGSDNSKAYVSGIINVDGAKFSESDNGVRIKTVQGGSGTAKNIKFNIRMEVKN 347

Qy 354 PIIDQNYCDRVPCIQOFSAVQVKNVYENIKGTSATKVAIKFDCSTNPFCEGIIEMENI 413  
Db 348 PIIDDDYDCK-DKCDQDSASAVQVKNVYENIKGTSATDVAITLNCSEKKPCQGIYLENV 406  
Qy 414 NLVBSGKPESEATCKNVHFNNAEHVTPHCT 443  
Db 407 KIRG-----GTASCKNANVNGQTVSPKCS 431

## RESULT 5

H84846  
probable polygalacturonase [imported] - Arabidopsis thaliana  
C/Species: Arabidopsis thaliana (mouse-ear cress)  
C/Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 09-Jul-2004  
C/Accession: H84846  
R/Lin, X.; Kaul, S.; Rounale, S.D.; Shea, T.P.; Benito, M.T.; Town, C.D.; Fujii, C.Y.;  
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanaken, S.E.; Umayam, L.; Tallon, L.;  
Euse, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.  
Nature 402, 761-768, 1999  
A/Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.  
A/Reference number: A84420; MIMD:20083487; PMID:10617197  
A/Accession: H84846  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-426 <STO>  
A/Cross-references: UNIPROT:O22935; GB:AE002093; NID:G2335094; PIDN:AAC02763.1; GSPDB:GNC  
A/Genetic: A2941850  
A/Map position: 2  
C/Superfamily: polygalacturonase

## Query Match

Best Local Similarity 49.8%; Score 1115.5; DB 2; Length 426;  
Matches 222; Conservative 74; Mismatches 117; Indels 33; Gaps 7;

Qy 6 NSILLIIIFASSISTCRSNVIDNLEFKOY-----DNLEQFAHDFQAYLSYLSKITE 60  
Db 6 NLVTFLMALMLFSGCKASRISPNVYDHSYKFKSDSLIKR--EDITGRSFFVASLR 63  
Qy 61 SNNNIDKVDKNGIKYINLVSPGAKGDKTYDNIAEFOAMNEACSSRTPOFVVPKKNYL 120  
Db 64 TPT-----TVSVDFGAKGDKTDQGAFFNMAWKACSSNGAVNLVLPKNTYL 112  
Qy 121 LKQITSPCRSSISVKIFGSLSEASSKISDYKDRRLMIAFDSVNLVVGCG--GTINGNG 178  
Db 113 LKSIQTLGPNLSILYQIFGTLASQKRSYKDIKIMFEDGNMNSVDSGDTGVVDGNG 172  
Qy 179 QVWPPSSCKINSLPCRDAPTLTFMNCCKNLKNNLKSNAQOHHIKESCTNVVASNLM 238  
Db 173 ETWQNSCKRNKA-----KALTFFNSKSLYKNLKNVNAQOIQISIECKSNVQSNV 225  
Qy 239 INASAKSPNTDGVHVSNTQYIQISDTIIGTGDDCISIVSGSNVQATNITCGPGHGISIG 298  
Db 226 VTAPADSPMTDGIHTNTQNIKRVNSDIDGTGDCISIESGQNVQINDITCGPGHGISIG 285  
Qy 299 SLGSGNSEAVYNSVTNVEAKITGAENGVRIRKTVQGGSGOASNIKFLVEMODVKYPIID 356  
Db 286 SLGDDNSKAFVSSVTVYDGAHLSGTDNGVRIRKTVQGGSGTASNIIFONIDNVKNPIID 345  
Qy 359 QNYCDRVPCIQOFSAVQVKNVYENIKGTSATKVAIKFDCSTNPFCEGIIEMENILVGE 418  
Db 346 QDYCDK-SKCTTEKSAVQVKNVYRISGTSASENNITNCSKNYCCGIVLDRVNIKG-- 403  
Qy 419 GSKPSEATCKNVHFNNAEHVTPHCTS 444  
Db 404 ---GRATCTNANVVDKGAVALPQCN 425

## RESULT 6

S71523  
polygalacturonase (EC 3.2.1.15) [similarity] - peach  
N:Alternate names: endopolygalacturonase

RESULT 7

polygalacturonase (EC 3.2.1.15) - apple tree  
N:Alternate names: poly [1,4-a-D-galacturonide] glycan hydrolase  
C:Species: Malus domestica (apple tree)  
C:Date: 15-Oct-1999 #sequence\_rev 15-Oct-1999 #text\_change 09-Jul-2004  
R:Accession: U17011  
R:Author: R. G. Plant Physiol. 105, 1437-1438, 1994  
A:Title: A cDNA clone for endopolygalacturonase from apple.  
A:Reference number: Z16649; PMID:95062722; PMID:7972500  
A:Accession: U17011  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-460 <ATK>  
A:Cross-references: UNIPROT:P48978; EMBL:U27743; NID:g456091; PIDN:AAA74452.1; PID:g45606;  
A:Experimental source: strain Golden delicious; ripe fruit  
C:Function:  
A:Description: catalyzes hydrolysis of 1,4-alpha-D-galactosiduronic linkages in galacturonic  
A:Pathway: polyaccharide degradation  
C:Superfamily: polygalacturonase  
C:Keywords: glycosidase; hydrolase; polysaccharide degradation

Query Match 45.6%; Score 1091; DB 2; Length 460;  
Best Local Similarity 49.3%; Pred. No. 2.9e-70;  
Matches 231; Conservative 76; Mismatches 126; Indels 36; Gaps 12;

QY 1 MWIQRSNLLILITFASISIT--CRSNVIDD-MLFKQVYNNILEQEPFADHFOAYLSTLSK 57  
DB 1 MALKTQLMSFVVFVVFVFSFTSCSGSFQEVNALHSYVDHVDKESGVNRAVSTY-TD 59  
QY 58 NIE-----SNNNIDKVDKNGI-----KYNLVSRGAKDGKTYDNIAF 95  
DB 60 TIEGLAKWELIRPTQLFSSRKANTI-TGSIATSSAPAKTISVDDFGAKNGAD-DIOAF 117  
QY 96 EOMANEACSSRTVQVVPFNKNVYLKQITFGSPCRSSISVKKIFGSLSEASSKISDPKDR 155  
DB 118 VKMKRAKACSSGAMVLVWPQ-KNYLVPRLEFSGPCSKQLTQIGRIEASBDSIYKDD 176  
QY 156 LMTAFDSVQVLVVGGGGTINGNGQVWVPSSCKTNKSLPERD-APFTALTWNCNKLKNNL 214  
DB 177 HWLIFDNVQNLVVGSGTITNGNGNTWKNKSCIKKPOPPGTYAPAVATYNNRCNNLVKNL 236  
QY 215 KSKNAQOHIKESECTNNVVASNMIMNASAKSNPTDGVHNSNTQYQIISPTITGTPDDCS 274  
DB 237 NIQDAQCHVHITONCINQVQASCLTWTAPEDSPNTDGIHVTMNTITSSVIGTDDCS 296  
QY 275 IVSGSONVATNITGCPHGHSISGLSGSNSFAVYNTVNEAKIIGABNGVRIITWOGG 334  
DB 297 IVSGSQRVQATDITGCPHGHSISGLSGEDGSDHVSQVFNVGAKLSGTNSNGIRITWKG 356  
QY 335 SGQANIEFLAVNEMDVYKPIIITIDNYCD-RVEPCIQQFSAVQVKNVYENIKGTSATKV 393  
DB 357 SGSATNIVQVQVQNDVNPPIIIDNYCDHKTCKCKQKQSAVQVKNVLYQNIIRITSASGD 416  
QY 394 AIKPDGCTNPFCEGIEMENINLVGSGKPSRATKKNHFNNAEHVTPHC 442  
DB 417 AITLNCQSVPCQGIIVLQSVQL-QNGR--AECCNVQPAVKGVVSPRC 460

RESULT 8

D96833  
hypochemical protein F18B13.25 [imported] - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 02-Mar-2001 #sequence\_rev 02-Mar-2001 #text\_change 09-Jul-2004  
R:Accession: D96833  
R:Theologos, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,  
Chen, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;  
Nansen, N.F.; Hughes, B.; Hultzer, L.  
Nature 408, 816-820, 2000





```

Db      107 VFRGFCQPHLSFKVNDGTIAQDPDPAWKNKSKIMLOFAQIUDTNFMIGTGVIDGQGOQWMAQ 166
QY      185 SCK-INKSLPC--RDAPFALTFMNCNKLKVNNLKSKNAQOIHIFESCTNVVASNLMINA 241
Db      167 QCKVNVGRTVCNDRNRFPALIKIDYSKSTVAKELTLMISPEFHLVFGCEGKIGLKLTKA 226
QY      242 SAKSPNTDGVHYNSNTQYQISDTIIGTDDCISIVSGSNVQANNTICGCGHGISISLG 301
Db      227 PRDSPNTDGDIDIPASKRPHIEKVICTDDCIAGTSSNITTKIDLCGCGHGISISLG 286
QY      302 SGNSEAVYNTVNEAKIIGANGVRIKTMQSSGSAKNTKFLNVEQDVKYPIIIDONY 361
Db      287 RDNRAEYSHVHVRNAPFIDTQNGLRITKQSSGSLASYITENVENINSENPILNQFY 346
QY      362 CDRVEPCIOQPSAVQVKNVYENIKGTSATKVAIKFPCSNPCEGIMENINLVGSGK 421
Db      347 CTSASACONORSAVQIOGVYIKNHIGTSATPAALIQMCSDSVPTGSIQLSNVLKLTSGK 406
QY      422 PSEATCKNVHFNNAEHVTPHCTSL 445
Db      407 PASCVDMNARGFYSGRLIPTCKNL 430

```

## RESULT 11

Probable polygalacturonase F2SP12.85 [imported] - Arabidopsis thaliana  
 C:Species: Arabidopsis thaliana (mouse-ear cress)  
 C:Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 09-Jul-2004  
 C:Accession: A96609  
 R:Theologas, A.; Ecker, J.R.; Palm, C.D.; Federpiel, N.A.; Kaul, S.; White, O.; Alonso, J.; Chiu, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.; Hansen, N.F.; Hughes, B.; Huizar, L.  
 Nature 408, 816-820, 2000  
 A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.; Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
 A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, K.; Wei, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
 A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
 A:Reference number: A66141; MUID:21016719; PMID:11130712  
 A:Accession: A96609  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-434 <STO>  
 A:Cross-references: UNIPROT:Q9FXC1; GB:AE005173; NID:g9954742; PIDN:AAG09093.1; GSPDB:GN  
 A:Gene: F2SP12.85  
 A:Map position: 1  
 C:Superfamily: polygalacturonase

```

Query Match      32.9%; Score 787.5; DB 2; Length 434;
Best Local Similarity 41.3%; Pred. No. 1,2e-48;
Matches 170; Conservative 68; Mismatches 155; Indels 19; Gaps 8;

QY      54 YLSKNISNNNDIVDKNGIKVINVSFGAKGCKTYDNIAPFOANNEACSSRTP--VOF 111
Db      26 YLSSPAPNPNAYNDNDIAPTFVDVTSFGAIGCSTDDTSAFKMAWDAACMSTGKSKLL 85
QY      112 VPRKNKYLKQITFGSGCRSSISVKIFGSLA---SSKISDYDRRLMIFSDVQULV 167
Db      86 LVPTFCFLVKKPTTFNCPCTNLVLIIDGFIVSPGPRSPSNY--QKQMMVFYVNGLS 143
QY      168 VGGGTTINGGVWMPSSCKINKSL-----PCRDAPFALTFFNCKNLKVNNLKSKNAQ 220
Db      144 IQSSGVINGGQKMWMLPCKPHKGLNGTQTGPFC--DSPVALRLFGSSKVRLOGINPMNSA 202
QY      221 QIHKESECTNVVASNLMINASAKSPNTDGVHVSNTQYQISDTIIGTDDCISIVSGSQ 280
Db      203 QEFVRFDNCSDVVVSIIKAPASSPNTDGIHENTHNVQIRNSVTSNGDDCISIGAGCF 262
QY      281 NVQANTITGCPGHGISISLGSGNSEAVYNTVNEAKIIGANGVRIKTMQSSGSAQN 340

```

```

Db      263 NVDIKNVTCGPRSHGISISLGVNNSQAVYSNITVTNSTWNSDNGVRIKTMQSSGSSVR 322
QY      341 IKFLANEMQDVKYPIIIDONYCDRVEPCIOQPSAVQVKNVYENIKGT-SATKVAIKFPC 399
Db      323 IVFSNITLWNVNRNIMIDYYCC-QTNNCANQTSVATSIISDLVYNIGITDLSAPPIHFGC 381
QY      400 STNPFCEGIIEMENINLVGSGKPSK-ATCKNVHFNNAEHVTPHCTSLSEI 450
Db      382 SDSVPCINLTLTBVDLFPSSKQHLNPFCKMNAVGSMTITVPVYCLLDAPPD 433

```

## RESULT 12

Cry j II protein - Japanese cedar  
 S48730  
 C:Species: Cryptomeria japonica (Japanese cedar)  
 C:Date: 07-May-1995 #sequence\_revision 21-Jul-1995 #text\_change 09-Jul-2004  
 C:Accession: S48730  
 R:Namba, M.; Kurose, M.; Torijoe, K.; Hino, K.; Taniguchi, Y.; Fukuda, S.; Usui, M.; Kuri  
 FEBS Lett. 353, 124-128, 1994  
 A:Title: Molecular cloning of the second major allergen, Cry j II, from Japanese cedar pc  
 A:Reference number: S48730; MUID:95010777; PMID:7926035  
 A:Accession: S48730  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-514 <NAM>  
 A:Cross-references: UNIPROT:P43212; GB:D37765; NID:g577695; PIDN:BA07021.1; PID:d100759f

```

Query Match      32.9%; Score 787.5; DB 2; Length 514;
Best Local Similarity 38.5%; Pred. No. 1,6e-48;
Matches 173; Conservative 78; Mismatches 163; Indels 35; Gaps 9;

QY      8 ILLIITIPASSISTKRSNVIDNLFKQYDNIIEQEFHAFQAVLYLSKNISSNNINDK 67
Db      12 VMQQLIMAAA-----EDQSAQIMLSDIEQ-----YL-----RSNNSLRK 47
QY      68 VD---KNGIKYINLSFGAKGDKTYDNIAPFOANNEACSSRTPVQFVVPKKNYLLKQI 124
Db      48 VHSRHDALINIFNVKXGAVGDGHDCTEAFSTWQAAC-KKPSAMLVFGNKKFPVNNVL 106
QY      125 TFSGCRSSISVKIFGSLAASSKISDYKRRRLMIAFDSVQNLVVGGGTTINGGVWMP 184
Db      107 FFGNPGCPHFPTFKVDGIIAAVONPASWKNRIMLOFAKLITGLFMKGVIIDGQKQWMAQ 166
QY      185 SCK-INKSLPC--RDAPFALTFMNCNKLKVNNLKSKNAQOIHIFESCTNVVASNLMINA 241
Db      167 QCKVNVNREICNDRPRTAIKFPSTGLIIQGLKLNISPEFHLVFQNCGGVYIIIGISITR 226
QY      242 SAKSPNTDGVHYNSNTQYQISDTIIGTDDCISIVSGSNVQANNTICGCGHGISISLG 301
Db      227 PRDSPNTDGDIDIPASKRPHIEKVICTDDCIAGTSSNITTKIDLCGCGHGISISLG 286
QY      302 SGNSEAVYNTVNEAKIIGANGVRIKTMQSSGSAKNTKFLNVEQDVKYPIIIDONY 361
Db      287 RENSRAEYSHVHNGAKFIPTQNGLRITKQSSGSLASYITENVENINSENPILNQFY 346
QY      362 CDRVEPCIOQPSAVQVKNVYENIKGTSATKVAIKFPCSNPCEGIMENINLVGSGK 421
Db      347 CTSASACONORSAVQIOGVYIKNHIGTSATPAALIQMCSDSVPTGSIQLSNVLKLTSGK 406
QY      422 PSEATCKNVHFNNAEHVTPHCTSL 448
Db      407 --IASCLNDNANGFYSGHVIAPACKNLSPS 433

```

## RESULT 13

Second major allergen Cry j II precursor - Japanese cedar  
 JC2498  
 C:Species: Cryptomeria japonica (Japanese cedar)  
 C:Date: 16-Mar-1995 #sequence\_revision 26-May-1995 #text\_change 09-Jul-2004  
 C:Accession: JC2498; PC2346; A60147  
 R:Komiyama, N.; Sone, T.; Shimizu, K.; Morikubo, K.; Kino, K.  
 Biochem. Biophys. Res. Commun. 201, 1021-1028, 1994  
 A:Title: cDNA cloning and expression of Cry j II, the second major allergen of Japanese c



```

Db      827 LEFSKLGKGVFOGKGVIDGSGKWMWASCKKKS-----NALTIESSGKVSGLTIQ 879
QY      218 NAOIHIKFESECTNVVNSMLMINSASPNITGVHVSNTQYIOISPTIIGTGDDCISIVS 277
Db      880 NSQOMNFIARSDSVRVSKVWSSPGDSPNTDGHITGSTNVLIDCKIGTGDDCVSIVN 939
QY      278 GSONVQATNITCGPBGHISIGSLGSGNSEAYVSNVTVNEAKIIGAENGVRITKW----- 331
Db      940 ASSNKKMKNITCGPBGHISIGSLGKDNITGIVTQVLDYALRETTNGLRITKYOVHKK 999
QY      332 -----OQSGQASNIKELVNEMQDVKYPPIIDONVCDRYE 366
Db      1000 KASLFSKNFPLASTTILFHCFSEQSGSGYVGIRFTVWEMQDVANPILIDQFYCDSP 1059
QY      367 PCTIOQBSAVOVKNVVENIKGTSATKVAIKFDCSTNFPCEGIIMENTNLVGSSEKPSSEAT 426
Db      1060 TCONQTSVAVKISQIMYRNITGTTKSAKAIKFACSDTVFCSHIVLNNVNLGNDQ-VEAY 1118
QY      427 CKNVHFNNAH-----VTPHCTSLISEDEAL 453
Db      1119 C-----NSAEGFGYVHHSADCLYSHDKGL 1145

```

Search completed: March 24, 2005, 07:05:59  
 Job time : 43 secs